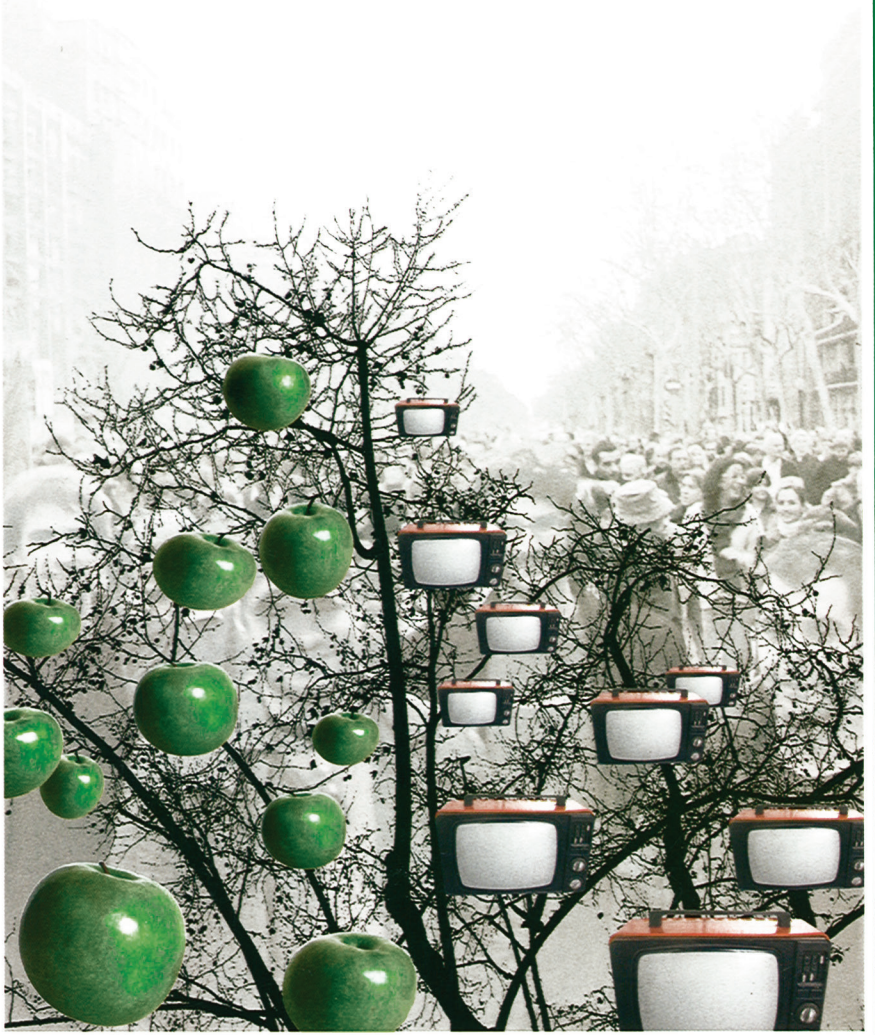
Filosofía/Sociología

Jon Elster

Lógica y sociedad

Contradicciones y mundos posibles



gedisa
editorial

Jon Elster

LOGICA Y SOCIEDAD

Grupo: CIENCIAS SOCIALES
Subgrupo: FILOSOFÍA

Editorial Gedisa ofrece
los siguientes títulos sobre

FILOSOFIA

pertenecientes a sus diferentes
colecciones y series
(Grupo "Ciencias Sociales")

JON ELSTER	*Lógica y sociedad*
MARTIN HEIDEGGER	*Introducción a la metafísica*
GIANNI VATTIMO	*La secularización de la filosofía. Hermenéutica y posmodernidad*
ROBERT NOZICK	*Meditaciones sobre la vida*
JON ELSTER	*Juicios salomónicos*
IAN HACKING	*La domesticación del azar*
THEDOR VIEHWEG	*Tópica y filosofía del derecho*
GEORGE STEINER	*En el castillo de Barba Azul*
PIERRE GRIMAL	*Los extravíos de la libertad*
JON ELSTER	*Tuercas y tornillos. Una introducción a los conceptos básicos de las ciencias sociales*
E. BARBIER, G. DELEUZE Y OTROS	*Michel Foucault, filósofo*
JOSÉ MARÍA BENEYTO	*Apocalipsis de la modernidad*
GREGORIO KAMINSKY	*Spinoza: la política de las pasiones*
PIER ALDO ROVATTI	*Como la luz tenue*
GEORGES BALANDIER	*El desorden*

(sigue en la pág. 287)

LOGICA Y SOCIEDAD

Contradicciones
y mundos posibles

por

Jon Elster

gedisa
editorial

Título del original en inglés:
Logic and Society. Contradictios and Possible Worlds
© 1978 by John Wiley & Sons, Great Britain

Traducción: Margarita N. Mizraji
Revisión técnica: Cecilia Hidalgo

Diseño de cubierta: Roberto Suárez

1ª reimpresión: abril del 2006, Barcelona
Reimpresión, 2019

Derechos reservados para todas las ediciones en castellano

© Editorial Gedisa, S. A.
Av. Tibidabo, 12, 3o.
08022, Barcelona, España
Tel. 93 253 09 04
gedisa@gedisa.com
www.gedisa.com

eISBN: 978-84-18193-82-8

Queda prohibida la reproducción total o parcial por cualquier medio
de impresión, en forma idéntica, extractada o modificada,
en castellano o en cualquier otro idioma.

Indice

Reconocimientos	9
Prefacio a la edición reimpresa	11
Notas	25
Introducción: Lógica y sociedad	29
Notas	38
1. La lógica formal. Una exposición informal	39
Apéndice: Lógica relacional y relaciones sociales	50
Notas	56
2. Razonamiento posibilístico en las ciencias sociales	59
Movilidad social	59
Gramáticas de estados finitos	62
Los sistemas de parentesco y el pensamiento estructuralista	64
Evaluación del ingreso nacional neto	67
Cambios endógenos de los gustos	71
Apéndice: Un resultado de imposibilidad para gramáticas independientes del contexto	75
Notas	77
3. Posibilidad política	81
Notas	97
4. Contradicciones de la mente	100
La teoría de las contradicciones de Hegel	102
Deseos contradictorios: el amo y el esclavo	106
Deseos contradictorios: la economía de la irracionalidad	113
Creencias contradictorias: Hintikka	118
Creencias contradictorias: Festinger	124
Notas	128

5. Contradicciones de la sociedad 133
La falacia de composición 135
Contrafinalidad 144
Suboptimalidad 162
Una teoría dual del cambio social 175
Apéndice 1: Algunas nociones y problemas de la teoría de juegos 194
Apéndice 2: Causalidad e intencionalidad: tres modelos de hombre 201
Notas 207

6. Los contrafácticos y la nueva historia económica 219
Tres capítulos sobre la historia del concepto 220
El problema de los contrafácticos históricos 225
Imperialismo y colonialismo 238
Optimismo y pesimismo 242
Las Leyes de Navegación 248
Los ferrocarriles y el crecimiento económico norteamericano 251
Los esclavos norteamericanos y su historia 257
Notas 268

REFERENCIAS BIBLIOGRÁFICAS 272
INDICE TEMÁTICO 284

Reconocimientos

Muchos amigos y colegas han contribuido durante años a las ideas que se encuentran en este libro. Me gustaría agradecer, ante todo, a Raymond Aron, quien me enseñó que la sociología y la economía son disciplinas históricas. Igualmente, a Per Schreiner y Dagfinn Føllesdal: ellos me transmitieron con el ejemplo y la crítica lo que la honestidad intelectual significa y requiere. También me he beneficiado en múltiples discusiones con Hans Fredrik Dahl, Gudmund Hernes, Knut Midgaard y Francis Sejersted. Los borradores fueron vistos por las siguientes personas: Brian Barry, Gudmund Hernes y Arthur Stinchcombe, quienes leyeron todo el texto y hallaron tiempo para entregarme comentarios detallados por escrito. Sin su ayuda, el libro habría estado más pobremente organizado y más atiborrado de digresiones confusas. Debo agradecimientos especiales a Brian Barry, por animarme a desarrollar el trabajo con la extensión de un libro; a Amélie Rorty y Ludwig Siep, por leer y comentar el capítulo 4; a Raymond Boudon y Michael Taylor, por sus opiniones acerca del capítulo 5; a Dagfinn Føllesdal, por salvar enojosos errores en el capítulo 6. He contado también con la invalorable ayuda de Aanund Hylland, quien me señaló errores conceptuales y matemáticos en los capítulos 1, 2 y 5. Sus contribuciones son reconocidas en los lugares correspondientes.

J.E.

Prefacio a la edición reimpresa

Desde que apareció la primera edición de *Lógica y sociedad* en 1978, tanto la lógica como las ciencias sociales han conocido importantes progresos. Mi dominio de la lógica formal era bastante rudimentario, incluso para el nivel alcanzado en 1978 y no he tratado de mantenerme al tanto de los avances más recientes. Por eso, era impensable una actualización del libro en ese sentido. Estoy mucho más al tanto de los progresos en las ciencias sociales. Pero en este terreno mis puntos de vista han cambiado en varios sentidos, por lo cual mi visión de las cosas es distinta de la de doce años atrás. Reescribir el libro para incorporar esos avances hubiera sido una tarea ciclópea, divergente de mis intereses del momento. Por eso resolví que se reeditara el libro con algunas modificaciones menores, principalmente para corregir errores de imprenta. No obstante, quiero referirme a ciertos avances recientes y a mis propios cambios de opinión sobre diversos asuntos. Se recomienda al lector que aborda este libro por primera vez que lea el texto original antes de pasar al resto de este prefacio, en el que se analizan diversas objeciones, agregados y eliminaciones aparentemente justificados por los nuevos avances. Me referiré a cada capítulo, a partir del tercero.[1]

Capítulo 3: Posibilidad política

Para empezar, haremos algunos comentarios anecdóticos sobre el concepto de la negación y, más concretamente, la "negación de la negación" (págs. 94-95). Luego, en tono más serio, analizaré una objeción importante al concepto de posibilidad política.

Spinoza acuñó la frase "Toda afirmación es negación", que luego se convirtió en un principio importante de la filosofía de Hegel. Este subrayaba que toda negación debía ser determinada, una negación de algo. Para ilustrar el concepto veamos el siguiente ejemplo de un menú: "Emparedado con jamón: $2. Emparedado con queso: $1. Emparedado sin jamón: $0,50. Emparedado sin queso: $0,25". Este ejemplo es absurdo, pero no lo son

otros, que ilustran un concepto distinto de la negación. Como señalo más adelante, en la *negación interna* en particular importa el objeto de la negación.

Los chistes lacónicos de Sidney Morgenbesser son famosos en los círculos filosóficos. Es el autor de la "primera ley de la lógica judía": Si *p*, ¿por qué no *q*? En respuesta al profundo interrogante de Martin Heidegger, "¿Por qué hay algo en lugar de nada?", respondió, siempre en el espíritu judío: "Si no hubiera nada, igual te quejarías".[2] En una conferencia alguien explicó con mucho palabrerío que, si bien una doble negación equivale a una afirmación, una doble afirmación jamás expresa una negación. Al escucharlo, Morgenbesser exclamó desde la platea: "¡Sí, sí!"

No he podido descubrir el origen de la siguiente anécdota, de manera que no garantizo que sea precisa en todos los detalles, pero creo que la esencia sí lo es. En la década de 1920, un miembro famoso del movimiento judío mundial sentía enormes celos de la prominencia creciente de Freud. Asimismo, lógicamente, demostraba gran escepticismo por la teoría freudiana del inconsciente. Por fin, entrevistó a Freud en Viena y le preguntó:

–En su opinión, doctor Freud, ¿quién es hoy la personalidad judía más destacada del mundo?

–Diría que es usted, doctor X.

–No, no –respondió éste con modestia.

–Un "no" hubiera bastado –observó Freud.

En efecto, un rechazo exagerado puede ser una señal de aceptación. En lógica, "no *p* y no *p*" equivale a "no *p*". Psicológicamente bien puede significar "no (no *p*)", por consiguiente, "*p*".

Considérese, por último, el intento infructuoso de Marx de demostrar que la operación matemática de la diferenciación es un caso de negación de la negación. En primer lugar, observa que si en la función $y = ax$ aumentamos x a x_1, y por consiguiente y a $y_1 = ax_1$, la diferencia $y_1 - y = a(x_1 - x)$ se reduce a $0 = 0$ cuando x_1 tiende a x. A continuación observa que "postular la diferenciación y luego suprimirla conduce literalmente a *nada*. La dificultad para comprender la operación de diferenciación —y *la negación de la negación en general*— radica precisamente en demostrar cómo difiere de este procedimiento sencillo y por lo tanto conduce a resultados verdaderos."[3] Este, como otros ejemplos del pensamiento dialéctico, me recuerda el truco mágico de la moneda que desaparecía, que una vez vi por televisión. El mago puso una taza sobre una moneda y aseguró que podía desplazarla a otra taza con un pase de manos. Agitó las manos en el aire y dijo: "Ahora la moneda pasó a la otra taza." Las agitó nuevamente y dijo: "Ahora está nuevamente debajo de la primera taza" y procedió a demostrarlo.

Hablando con seriedad, se trata de saber si la lógica modal puede representar apropiadamente la posibilidad política. Anthony Kenny[4] rechaza ese postulado en el caso de la capacidad. Objeta[5] que la ley de distribución de posibilidades sobre la disyunción no es aplicable a la capacidad. Supóngase que uno es capaz de acertar en el sector superior o inferior del blanco, es decir, de dar en el blanco. Ese puede ser el límite de su capacidad. Tal

vez sea falso que uno es capaz de acertar en la mitad superior o inferior. Pero la expresión de la capacidad como posibilidad debería satisfacer la ley distributiva, puesto que "si una disyunción es verdadera en un mundo posible, entonces uno de los disyuntos debe ser verdadero en algún mundo posible".[6]

Esta objeción se aplica a mi análisis de la posibilidad política. Es casi un axioma de la política que se puede provocar un estado sólo si queda incompleta la descripción de algún aspecto crucial. "Dejamos que el futuro elabore este problema."[7] (Por otra parte, tal vez sea verdad, como dice Marx, que para tener éxito en la acción política uno debe "resolver tanto los *detalles* como los principios antes de llegar al poder».[8] Para conquistar el poder, tal vez uno deba referirse vagamente a los detalles en la discusión pública; para conservarlo, uno debe elaborarlos previamente en la intimidad.[9]) Aunque todos saben que después de la revolución se deben tomar decisiones difíciles, podría ser políticamente imposible llevar a cabo una revolución si se plantea el problema y se lo resuelve de una u otra manera antes de la revolución. Esta verdad de la vida política es incompatible con la ley modal de la distribución de la posibilidad sobre la disyunción. Supóngase la posibilidad de provocar un estado en el que los soviets o el Comité Central detentan el poder. Si rigiera la ley, debería ser políticamente posible provocar un estado en el que el soviet tiene el poder o bien uno en el cual el Comité Central tiene el poder. Como se ha señalado, esto puede ser falso, de manera que la ley fracasa.

Una forma de resolver este problema sería la de limitarse a la posibilidad política cuando todos los actores son racionales.[10] La ambigüedad tal vez facilita el acuerdo porque todas las partes confunden sus deseos con la realidad y creen que el problema se resolverá de acuerdo con sus preferencias. Lamentablemente, este postulado es incorrecto. La política suele obligar a los actores a tomar decisiones en medio de la incertidumbre, un estado cognoscitivo en el cual es imposible formar creencias correctas o atribuir las probabilidades correspondientes a los distintos desenlaces posibles. Si no fuera por la incertidumbre, el concepto de posibilidad política perdería casi todo su interés. En esos casos se puede llegar al consenso porque los actores utilizan la heurística para tomar decisiones bajo la incertidumbre que los obliga a adoptar creencias mutuamente incompatibles del tipo "como si". Todos suelen actuar sobre la base de las presunciones más optimistas, que en situaciones de incertidumbre son racionalmente lícitas, pero no las más convenientes.

A esta altura no veo cómo se puede refutar esta objeción. Al mismo tiempo, no me parece que destruya totalmente el proyecto. El modelo de la posibilidad política es, como todos, imperfecto y en ocasiones deficiente. Los axiomas de la lógica modal abarcan una gran cantidad de percepciones sobre la posibilidad política, pero no todas, como se acaba de advertir. Tal vez valdría la pena internarse en los arcanos de la lógica modal en busca de un modelo más adecuado, pero sólo si cada una de las relaciones constitutivas de ese modelo tuviera un sentido político propio.[11] Pero a pesar de sus

deficiencias, el modelo sencillo esbozado en el capítulo 3 puede servir para centrar las ideas si –como considero que sucede en este caso– falla por razones capaces de orientarnos en direcciones que de otro modo no hubiésemos emprendido.

Capítulo 4: Contradicciones de la mente

Desde que escribí este capítulo he adquirido conciencia de otras contradicciones psicológicas. Una es el intento de *generar* intencionalmente estados que en el mejor de los casos sólo pueden surgir como subproductos de acciones emprendidas con otros propósitos. La otra resulta de confundir la *negación externa* de una proposición con su *negación interna*.

Puesto que me he explayado sobre el asunto en una obra ampliamente difundida,[12] aquí me limitaré a exponer las ideas fundamentales en forma sintética. Sostengo —con diversas modificaciones que no expondré— que ciertos estados *son esencialmente subproductos*, en el sentido de que el intento de provocarlos intencionalmente está condenado a fracasar. Los estados unipersonales paradigmáticos que una persona no puede inducir en sí misma a voluntad incluyen el sueño, la espontaneidad, la dignidad, la inocencia, la ironía romántica, la salvación, la fe, el amor, el olvido, la excitación sexual, al ser cosquilloso, el tener "carácter", la autoestima, hablar sin tartamudear. Los estados que uno no puede inducir intencional y públicamente[13] en otros incluyen la admiración, la gratitud, el amor, el respeto, la autoestima, el odio, la autonomía. Me extenderé brevemente sobre la autoestima, incluida en las dos categorías. A pesar de la amplia difusión de los manuales de autoayuda sobre la manera de inculcarse autoestima, existen buenos motivos para pensar que nadie puede promoverse a sí mismo por este medio.[14] Las autoridades públicas han intentado crear puestos de trabajo para los desempleados para ayudarlos a conservar su autoestima, pero han fracasado porque el trabajo no proporciona amor propio a menos que su producto principal posea valor independiente.[15]

Dada la naturaleza paradójica de los intentos de generar intencionalmente estados que la voluntad no puede crear, ¿de qué manera la lógica formal nos ayuda a comprenderlos? Creo que existe una analogía entre los criterios de consistencia para los intentos y los criterios de consistencia de Hintikka para las creencias. Según Hintikka, no basta decir que una creencia es consistente si y sólo si existe un mundo posible en el cual es verdadera. Debe existir un mundo en el cual sea verdadera *y creída*.[16] Si tratamos de comprender el intento de ser espontáneo o la orden de ser autónomo ("¡No seas tan obediente!") bajo esta luz, advertimos que el contenido de tales esfuerzos no tiene nada de autocontradictorio. Existen mundos en los que las personas se conducen de manera espontánea y autónoma. Pero no existe un mundo en el que la espontaneidad es a la vez realizada *e intentada* o la orden de ser autónomo es impartida *y obedecida*.

Me extenderé sobre la distinción entre la negación externa e interna.[17] Primero explico la fundamentación lógica de la distinción, luego me interno

en la historia del pensamiento para rastrear su desarrollo desde Kant hasta el presente y finalmente utilizo la obra del satírico disidente soviético Alexander Zinoviev para demostrar cómo puede aclarar las características de las sociedades reales. Sostengo que la confusión entre las dos formas de negación caracteriza una cierta actitud mental que por falta de un término mejor podemos llamar *la mentalidad primitiva*, dejando de lado las connotaciones peyorativas que la expresión podría haber tenido en el pasado.

Considérense las siguientes afirmaciones:

I. A cree *p*.

II. No (A cree *p*).

III. A cree No *p*.

Llamaré II a la negación externa de I, y III a su negación interna. La negación de la lógica formal es la negación externa. Así, en las "leyes del pensamiento" —el principio de contradicción y el del tercero excluido— la negación siempre tiene este sentido externo. El principio de contradicción dice "No (I y II)". El principio del tercero excluido dice "I o II". Por consiguiente, aceptar estos principios cuando se toma la negación en su sentido interno sería una manifestación de la mentalidad primitiva. La cita de Alastair MacIntyre en la página 118 ilustra esta confusión con respecto al principio de contradicción. La interpretación igualmente confusa del principio del tercero excluido aparece también cuando se confunde la apostasía con la herejía o el agnosticismo con el ateísmo. La encontramos en el "maniqueísmo cotidiano" de frases tales como "Quien no está conmigo está en contra de mí" (o, como dicen que dijo Janos Kadar en 1956, "Quien no está en contra de mí está conmigo") y "El enemigo de mi enemigo es mi amigo". Adviértase que en el modelo de Hintikka de la creencia racional, el principio de contradicción *es* válido para la negación interna, pero no así el del tercero excluido. El hombre racional no puede creer que Dios existe y a la vez que no existe, pero puede tener amplitud de miras y no adoptar una ni otra creencia.

Más adelante se verá cómo pueden aparecer confusiones similares bajo distintas interpretaciones del operador modal **N**. Pero aquí quiero exponer dos problemas formales estrechamente relacionados. El primero es la confusión entre la negación de una conjunción y la conjunción de las negaciones, que es característica de los niños en determinadas etapas de su desarrollo intelectual[18] y también se encuentra en otros contextos. En un relato folklórico noruego,[19] una niña buena y una niña mala deben superar una serie de obstáculos para obtener un premio. Para alcanzar el éxito final es necesario pasar cada una de las pruebas. De más está decir que la niña buena supera con éxito todas las pruebas y la mala *falla en todas*, aunque hubiera bastado un solo fracaso para quedar descalificada. Esta falacia permite en cierta medida comprender también la psicología de los estereotipos. En las sociedades tradicionales se concebía la excelencia como una e indivisible: quien se destacaba en un campo debía destacarse en todos; el fracaso en uno indicaba que fracasaría en todos.[20] Ahora bien, muchos

atributos están vinculados causalmente entre sí. Se suele suponer con fundamento que el poder acompaña a la riqueza. Pero, digan lo que dijeren lady Catherine de Bourgh y su obsecuente señor Collins, no nos permiten inferir la inteligencia o la belleza.

El segundo problema formal al que me refiero es la lógica de las descripciones determinadas. (Lo analicé brevemente en el capítulo 1, pero sin extenderme porque no tenía en mente aplicación alguna.) En un célebre artículo Bertrand Russell sostiene que una frase como "el rey de Francia" sólo tiene sentido en el contexto de una proposición.[21] Así, se puede analizar la afirmación "El rey de Francia es calvo" en el sentido de que existe un *x* tal que (i) *x* es rey de Francia, (ii) para todos los *y* tales que *y* es rey de Francia, *y*=*x*, y (iii) *x* es calvo. Esta afirmación hoy es falsa porque no existe un *x* tal que *x* es el rey de Francia. ¿Pero cómo se ha de entender la afirmación "El rey de Francia no es calvo"? Una interpretación externa de la negación conduce a que esta afirmación equivale a la negación de la conjunción de (i), (ii) y (iii). Interpretada internamente equivale a la afirmación de (i) y (ii) juntamente con la negación de (iii). En lenguaje cotidiano, la interpretación natural es la segunda. Cuando alguien afirma que "el rey de Francia no es calvo", no afirma la existencia de un rey en Francia, pero desde el punto de vista pragmático en realidad sí se presupone esa afirmación.[22] Asimismo, cuando uno le pregunta a otro si ha dejado de golpear a su esposa, no afirma que lo ha hecho en el pasado, pero se da por sentada esa afirmación. Un interrogador que exige una respuesta por Sí o por No muestra una conducta análoga a la persona que afirma el principio del tercero excluido con la negación en su sentido interno.

En un breve tratado precrítico, *Versuch den Begriff der negativen Grössen in die Weltweisheit einzuführen*, Kant aborda algunos de estos problemas. El contexto de la obra es una oscura controversia sobre la filosofía de la física que no nos interesa aquí.[23] A nuestros fines nos interesan principalmente los ejemplos que da Kant para demostrar la existencia de dos clases de negación. (i) La negación externa del movimiento es el reposo, la interna es el movimiento en la dirección contraria. (ii) La negación externa de la riqueza es la pobreza, la interna el endeudamiento. (iii) La negación externa del placer es la indiferencia o el equilibrio, ya que la primera corresponde a la ausencia de causas de dolor y placer, y el segundo a causas exactamente contrapuestas; la negación interna es el dolor. (iv) La negación externa de la virtud no es el pecado de omisión, el cual, no menos que el de obra, es una negación interna; sólo las deficiencias del santo y las faltas del noble constituyen negaciones externas de la virtud. (v) La negación externa de la atención es la indiferencia, la interna es la abstracción. Hoy diríamos que no es lo mismo la ausencia de conciencia de *x* que la conciencia de la ausencia de *x*. (vi) La negación externa de la obligación es el permiso, la negación interna es la prohibición. (vii) La negación externa del deseo es la indiferencia, la interna es la repugnancia. Hoy diríamos que no es lo mismo la ausencia de deseo por *x* que el deseo de ausencia de *x*.

Estos ejemplos revelan claramente que Kant confundía dos clases de fenómenos. Por un lado están los ejemplos (v), (vi) y (vii) que corresponden a la distinción modal entre N(No *p*) y No(N*p*). Por el otro lado, los ejemplos (i), (ii) y (iii) corresponden a la distinción matemática entre el elemento neutro y el elemento inverso en un grupo de transformaciones. En algunos casos las dos distinciones podrían coincidir. Así, Zinoviev sugiere que hay dos clases de personas en la sociedad soviética que obstaculizan la realización del bien. Algunas son extraordinariamente mediocres, es decir, muestran una notable ausencia de cualidades. Otras muestran una gama impresionante de cualidades negativas. Esto corresponde a la distinción matemática entre los elementos neutros e inversos. Pero también sugiere que los elementos neutros son indiferentes al bien, en tanto los inversos lo combaten activamente. Esto corresponde a la distinción modal.

Zinoviev aplica la distinción modal entre la negación externa e interna a una serie de fenómenos. El más notable es, tal vez, su observación de que el proceso de desestalinización se realizó de manera totalmente estalinista. Corresponde a la esencia del estalinismo que lo que no es obligatorio está prohibido. Después de la muerte de Stalin, sobrevino un breve período durante el cual dejó de ser obligatorio citar sus palabras. En poco tiempo, se hizo obligatorio no citarlo. Zinoviev también menciona el caso de un disidente que se negó a negarse a que lo postularan a la Academia de Ciencias, y todos interpretaron que su actitud era la de buscar activamente la postulación. Esta tergiversación del principio de que la negación reiterada equivale a una afirmación surge de confundir la negación interna con la externa, el No(No(N*p*)) con No(N(No *p*)). Así, las mismas relaciones internas entre los disidentes responden al principio fundamental del régimen al que se oponen. Esta distinción permite aclarar un dilema que enfrenta el régimen en sus relaciones con la oposición. ¿Debe ignorar a la oposición (negación externa) o reprimirla (interna)? La primera alternativa se confunde fácilmente con la debilidad. La segunda equivale a un reconocimiento implícito de la oposición y a veces ayuda a difundir sus ideas. Así, un libro que criticaba el arte occidental se agotó rápidamente debido a sus reproducciones de pinturas que ejemplificaban las tendencias decadentes.

El silencio mismo puede revelar la existencia de atención y equivaler a un reconocimiento implícito. Si uno es verdaderamente indiferente hacia alguien, puede mencionar su nombre de vez en cuando, cuando el contexto lo requiere. Si uno nunca lo menciona, esa actitud puede ser una señal de interés y odio obsesivos (la actitud del régimen soviético hacia los disidentes) o de amor (la actitud de una esposa infiel hacia su amante). En un caso más o menos similar, no se debe confundir el azar —la negación externa del orden— con la mezcla intencional, una negación interna que revela su orden intrínseco por la ausencia de los agrupamientos que caracterizan a los fenómenos verdaderamente aleatorios. Existen pruebas (aleatorias) de que la selección de preguntas de diversas subdisciplinas por parte de los examinadores universitarios en años sucesivos responde a una mezcla

intencional más que a una selección al azar. Teóricamente, los estudiantes deberían poder explotar esa práctica en provecho propio.

El siguiente informe de Pekín (*The International Herald Tribune* del 25 de julio de 1989) es un ejemplo excelente de la distinción entre la negación externa y la interna y la confusión que, según Zinoviev, caracteriza a los regímenes totalitarios:

> Un funcionario chino sonreía con orgullo y alivio al acompañar a una visitante extranjera en un paseo en auto alrededor de la plaza Tiananmen.
>
> —Observe qué normal está la plaza —dijo, volviéndose hacia su invitada en busca de un gesto de aprobación.
>
> La invitada, una señora norteamericana residente en Hong Kong, contempló la plaza desierta, vigilada por soldados armados.
>
> —¿Dónde está la gente? —preguntó—. Está desierta. Lo siento, no me parece muy normal.
>
> El funcionario reconoció con renuencia que la plaza, un lugar transitado por los turistas que en mayo se había transformado en un centro de la democracia, estaba cerrada al público.

Estos ejemplos —ignorar a la oposición, jamás mencionar al amante, mezclar intencionalmente las preguntas de los exámenes, cerrar la plaza Tiananmen— poseen una estructura común. La verdadera indiferencia no es sistemática. Los intentos de fingir o imitar la indiferencia suelen fracasar debido a su carácter excesivamente sistemático. Revelan la actitud de negación interna que subyace detrás de la apariencia deseada de negación externa.

Capítulo 5: Contradicciones de la sociedad

En este capítulo definí la falacia de composición como la inferencia de cualquiera a todos. Contra esta definición esgrimí el argumento de que "a la falacia en cuestión no corresponde una falacia de división, la proposición recíproca tradicional de la falacia de composición" (pág. 135) y agregué que "la recíproca de lo que llamo la falacia de composición es, en verdad, una inferencia válida" (*ibíd.*). La segunda afirmación es verdadera, la primera tal vez no. Aquí definiré una falacia que, si bien no es la recíproca exacta de la de composición, puede considerarse con fundamento la falacia recíproca.

La falacia de composición es la inferencia

$$(\forall x)(M(Fx)) \supset M(\forall x)(Fx) \qquad (1)$$

Defino la falacia de división así:

$$M(\forall x)(Fx) \supset (\forall x)M(\forall y)(Fy \supset (x=y)) \qquad (2)$$

Como se explica en la página 138, el miembro derecho de la afirmación (2) sostiene que para una *x* dada, es posible que *x* sea el único individuo que posee la propiedad F. Por consiguiente, la inferencia en su conjunto afirma que si se puede atribuir una propiedad simultánea y consistentemente a todos los individuos, también se la puede atribuir a cada uno de manera individual *y exclusiva*. En realidad no me referiré a la falacia así definida, sino a una inferencia un poco más general, en la cual el consecuente afirma la consistencia de atribuir la propiedad exclusivamente a cualquier subconjunto propio del conjunto de individuos (incluyendo, desde luego, los conjuntos unimembres). Si es posible que todos los cisnes sean negros, debe ser posible que uno (o diez) sea negro mientras todos los demás son blancos. Si es posible que los obreros de todas las empresas perciban salarios bajos, cualquier firma en particular podría ser la única que paga bajos salarios a sus obreros. En el ejemplo keynesiano paradigmático, la falacia de composición sostiene que si la propiedad de ser la única empresa que paga bajos salarios es válida para cualquier empresa, entonces puede serlo simultáneamente para todas. La falacia de división sostiene que si la propiedad de pagar bajos salarios puede ser válida simultáneamente para todas, entonces la propiedad de ser la única empresa en hacerlo puede ser válida para cualquiera de ellas.

Para que la inferencia (2) sea una falacia, la propiedad F debe ser de alguna manera *no individualizable*. En los ejemplos que se me han ocurrido, esta imposibilidad se debe a razones causales. (Recuérdese mi argumento de que la imposibilidad correspondiente en la falacia de composición puede deberse a razones lógicas, conceptuales o causales.) Concretamente, en muchos casos interviene la imposibilidad política. Empiezo con un ejemplo que incluye las falacias de composición y de división. En su análisis de la alibertad* proletaria, G.A. Cohen sostiene que la libertad de los trabajadores para abandonar la clase obrera a fin de emplearse por cuenta propia es un ejemplo de la falacia de composición.[24] Si postulamos que *cualquier* obrero puede hacerlo, esto no significa que *todos* pueden; por consiguiente, son colectivamente alibres, así como son individualmente libres. Si un obrero dado es libre para abandonar la clase obrera, es porque se puede esperar que la mayoría de los demás no traten de usar esa libertad. Si todos lo intentaran, no todos lo conseguirían. Recíprocamente, la razón por la cual la mayoría de los trabajadores no trata de abandonar la clase obrera se relacionaría con la falacia de división. Una situación en la que todos los obreros abandonaran las filas de la clase obrera podría resultar moralmente aceptable para los trabajadores. Esto establece el antecedente de (2) sobre la lectura deóntica de los operadores modales. Sin embargo, la solidaridad obrera no les permitiría aceptar una situación en la que algunos, pero no todos, abandonaran las filas de la clase. Así se demostraría la falsedad del consecuente en (2).

* He acuñado el neologismo *alibertad* (falta de libertad) como traducción del neologismo inglés *unfreedom* del original [T.].

De acuerdo con un viejo axioma, *Qui peut le plus, peut le moins*. Quien pueda levantar diez kilos, también puede levantar cinco. Pero en algunos casos el axioma falla porque incorpora la falacia de división. Así, un patrono podría licenciar a todos sus trabajadores, pero no a algunos en lugar de todos.[25] Con respecto al servicio militar obligatorio, Tocqueville observa que "un gobierno democrático puede hacer prácticamente lo que le plazca, siempre que sus órdenes rijan simultáneamente y para todos; es la desigualdad de la carga, no su peso, lo que suele provocar la resistencia". [26] Asimismo, sería políticamente posible ofrecer trasplantes de órganos a todos los que lo requirieran, pero no a un subconjunto propio de ese grupo. Desde ese punto de vista se comprende la moratoria aplicada recientemente en Oregon a los trasplantes no renales.[27] En ese caso, la intersección entre las medidas políticamente posibles y las económicamente factibles tenía un solo miembro: no ofrecer el bien escaso a nadie.

Capítulo 6: Los contrafácticos y la nueva historia económica

Haré dos comentarios sobre este capítulo, relacionados con lo que he llamado la paradoja básica de los contrafácticos. En primer lugar, daré un ejemplo de la *paradoja básica de los subjuntivos*. Luego analizaré el intento de Peter Morriss de eliminar la paradoja de los contrafácticos.

La paradoja básica de los contrafácticos es la siguiente. Supóngase que se desea saber qué hubiera ocurrido si se postulara que una parte del pasado fue distinta de lo que sucedió en la realidad. Para ello se necesita un conocimiento general o una teoría que permita efectuar una inferencia del antecedente hipotético al consecuente hipotético. Pero ese conocimiento en sí podría ser incompatible con el antecedente contrafáctico. Para afirmar contrafácticos con significado hay que establecer un equilibrio sutil: la teoría debe ser lo suficientemente consistente para permitir que se infiera el consecuente del antecedente, pero lo suficientemente débil para ser consistente con el consecuente.

La paradoja básica de los subjuntivos es la siguiente. El subjuntivo tiene la forma, "si sucediera X, se produciría Y". Nuevamente, se requiere algún conocimiento general o teórico para inferir Y de X, y también en este caso el conocimiento mismo podría invalidar todo el experimento intelectual. En ese caso, no es legítimo utilizar el subjuntivo para sustentar otras proposiciones.

El ejemplo que emplearé para ilustrar la paradoja de los subjuntivos ha surgido en discusiones recientes de la inducción retroactiva, un método de razonamiento derivado de la teoría del control óptimo que constituye una herramienta fundamental del análisis económico.[28] En lugar de definir la inducción retroactiva la ilustraré por medio de un ejemplo. Considérese el siguiente juego con dos participantes.[29] Es de *conocimiento común* que

los dos son racionales: cada uno lo es y sabe que el otro también lo es y sabe que el otro sabe que lo es y así sucesivamente. La naturaleza del juego es también de conocimiento común.

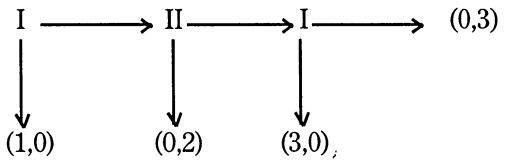


Figura 1

Juega uno por vez, empezando por I. Puede descender y poner fin al juego, en cuyo caso obtiene 1 y el jugador II nada, o avanzar hasta el segundo nodo y continuar. Entonces juega II, que puede descender, ganar 2 para sí y dejar nada para I o avanzar hasta el tercer nodo. Allí, I puede optar por ganar 3, sin dejar nada al otro, que tiene las opciones contrarias.

El desenlace parece evidente: I desciende en la primera jugada y pone fin a la partida. Para justificar esta conclusión se invoca tradicionalmente el principio de la inducción retroactiva, que razona desde el final del juego hacia atrás. Al comenzar el juego, I contempla lo que haría *si estuviera* en el último nodo del diagrama. Evidentemente, preferiría descender en lugar de avanzar. Puesto que sabe que II sabe que es racional, I anticipa que en el segundo nodo II descenderá para obtener 2 en lugar de 0, que es lo que obtendría si avanzara. Pero esa anticipación obliga a I a descender en la primera jugada para obtener 1 en lugar de 0, que es lo que obtendría si avanzara.

Este razonamiento parece irresistible. Pero contiene un problema: *¿por qué habría de contemplar I la posibilidad de ocupar el tercer nodo?* ¿Cómo habría de llegar, si los jugadores racionales siempre terminan el juego en el primer nodo? A los subjuntivos, como a los contrafácticos, se los puede afirmar cuando las premisas adicionales utilizadas para extraer conclusiones del antecedente hipotético son consistentes con el propio antecedente. No se puede afirmar que en ausencia del ferrocarril se hubiera producido igualmente el desarrollo de la economía porque se hubiera inventado el automóvil anteriormente, si esa afirmación se basa en una teoría de desarrollo de la técnica que permitiera también vaticinar la invención del ferrocarril. (Más adelante me extiendo sobre este ejemplo.) Tampoco se puede vaticinar que I descenderá si llega al tercer nodo sobre la base de la racionalidad de I, que no permite que el juego llegue hasta ese nodo.

Por consiguiente, se puede reformular el problema de la siguiente manera: ¿Existe un conjunto de presunciones sobre los jugadores lo suficientemente fuertes para inferir lo que harían en los distintos nodos del

juego y a la vez lo suficientemente débiles para ser consistentes con la posibilidad de llegar a esos nodos? He sostenido que la presunción de racionalidad y conocimiento común es demasiado fuerte para satisfacer la segunda condición. En casos especiales, otras presunciones pueden satisfacer las dos condiciones.[30]

(1) Primero, el jugador I podría presumir que se ha llegado a los nodos siguientes a causa de los errores cometidos (la presunción de la "mano temblorosa") . Es decir, se podría llegar al tercer nodo si existiera alguna probabilidad de que un jugador oprimiera el botón equivocado, por así decirlo. Si consideramos un juego similar al de la figura 1, pero de cien movidas sucesivas, la presunción se vuelve extremadamente improbable, ya que obligaría a cada jugador a cometer cincuenta errores no correlacionados antes de llegar al nodo final.

(2) Por consiguiente, la segunda presunción es que los errores sí son correlacionados, o sea que II es irracional. Desde el punto de vista de I, que se pregunta cómo pudo llegar al nodo final, es más sensato presuponer que II siempre avanza (la hipótesis del "autómata") y que ha llegado al nodo final aprovechando la irracionalidad de II.

(3) En tercer lugar, I podría contemplar una posibilidad más compleja, la de que II es un autómata irracional o bien un jugador racional que finge ser irracional a fin de inducir a I a avanzar en lugar de descender. En el juego ilustrado en la figura 1 esta presunción no tiene asidero, pero en el que se ilustra a continuación podría ser una explicación convincente de por qué se alcanzaron los nodos posteriores del diagrama.

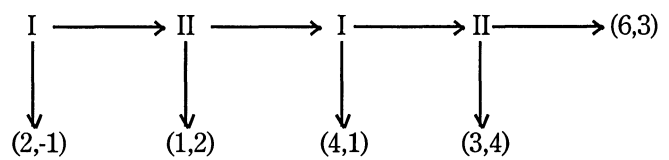


Figura 2

Para explicar por qué podría llegar al tercer nodo, I puede asignar la probabilidad p a que II es un autómata que sólo avanza en sentido horizontal y la probabilidad $1\ p$ a que II es un agente racional que finge la conducta de un autómata en las primeras etapas del juego a fin de inducir a I a avanzar para que él pueda descender más adelante.

(4) Por último, se puede debilitar la presunción de que la racionalidad de los jugadores es de conocimiento común.[31] En el juego ilustrado por la figura 1 sólo se presupone que I cree que II es racional, II cree que I es racional y I cree que II cree que I es racional. Cuando I contempla la posibilidad de encontrarse en el tercer nodo, debe preguntarse si la

presunción es consistente con el conjunto de creencias. Concretamente, si II hubiera podido avanzar en el segundo nodo. Para responder, antes debe explicar cómo II pudo racionalizar su presencia en el segundo nodo. En virtud de la estructura estipulada de creencias, no puede explicarlo con la presunción de que II cree que él es irracional, pero I puede suponer que II cree que él lo considera irracional, porque en ese caso II razonaría que I avanzó en el primer nodo con la esperanza de que él lo hará en el segundo. Ahora I sabe que si se llega al segundo nodo, II descenderá, porque cree que II cree que él es racional y por ende anticipará su descenso en el tercer nodo. Por consiguiente, desciende en el primero. El argumento de la inducción retroactiva es válido solamente cuando la racionalidad de los jugadores *no* es de conocimiento común. Si cada jugador posee inicialmente menos conocimientos sobre el otro, los dos pueden elaborar creencias estables. Con un conocimiento mayor, el subjuntivo hubiera sido imposible de aplicar.

Las presunciones (1) a (4) difieren entre sí de la siguiente manera. La presunción (1) estipula que I abriga cierta probabilidad subjetiva de que II es racional, pero falible. La presunción (2) estipula que I abriga cierta probabilidad subjetiva de que II es un autómata irracional. La presunción (3) estipula que si II es racional, será consciente de la evaluación de probabilidades hecha por I y tratará de explotarla en su provecho, y que I sabe que si II es racional será consciente de ello y así sucesivamente. La presunción (4) estipula que los jugadores son racionales e infalibles, pero su conocimiento de la racionalidad del otro y el conocimiento de ese conocimiento, etcétera, tiene un límite superior. En ciertos casos y bajo condiciones especiales, cada una de estas presunciones puede proporcionarnos lo que buscamos: una teoría lo suficientemente fuerte para permitirnos inferir qué harán los jugadores racionales en los distintos nodos y lo suficientemente débil para admitir la posibilidad de que lleguen a esos nodos. Pero ninguna de ellas parece poseer la simplicidad y la generalidad capaces de sustentar la inducción retroactiva en un contexto más general. En consecuencia, el principio en sí puede ser más vulnerable de lo que parece a primera vista. Con todo, sospecho que no se ha dicho la última palabra.

En su libro reciente sobre el poder, Peter Morriss sostiene que la paradoja de los contrafácticos no es *básica*, puesto que aparece solamente en el número limitado de casos en los que la *misma* teoría afecta tanto la legitimidad del antecedente como el vínculo entre éste y el consecuente.[32] Coincido en que la paradoja no aparece en todos los contrafácticos. Si esto significa o no que no es *básica*, es un problema semántico de escaso interés.

A continuación, Morriss sostiene que la paradoja no aparece en lo que yo cité como ejemplo paradigmático de su aparición, a saber, el estudio de Robert Fogel sobre los ferrocarriles y el desarrollo económico de Estados Unidos. Citaré su argumento in extenso:

> A Fogel le interesaba evaluar "la proposición de que los ferrocarriles eran indispensables para el desarrollo económico norteamericano durante el siglo

XIX" (Fogel, 1964, pág. vii), lo cual implica preguntarse, contrafácticamente, qué hubiera sucedido si no se hubieran desarrollado los ferrocarriles. Pero la versión de Elster del antecedente del contrafáctico –"si no se hubieran *inventado* los ferrocarriles"– es en sí un invento de Elster: Fogel jamás menciona la posibilidad de que los ferrocarriles no se desarrollaran porque no los inventaron. En cambio, menciona algunas razones por las cuales los ferrocarriles hubieran podido ser insignificantes –por ejemplo, que "las chispas lanzadas por las locomotoras al pasar podrían incendiar los edificios y los campos" (Fogel, 1964, pág. 2)–, pero eran todas razones esgrimidas por los primeros críticos del ferrocarril *después* de su invención.

Para usar el contrafáctico de Fogel como ejemplo de su paradoja, Elster se ve obligado a tergiversarlo de manera tal que su antecedente se refiera a la invención. Pero el verdadero contrafáctico de Fogel sólo contradice una creencia aislada (que los ferrocarriles se desarrollaron durante el siglo XIX), de manera que podemos suponer que el antecedente es verdadero. Si se nos pregunta cómo sucedió, basta que ejercitemos un poco la imaginación. Por ejemplo, podemos suponer que cuando inauguraron el ferrocarril, un cura dijo que era el artefacto del diablo; que el primer tren descarriló y ocasionó una gran mortandad; en consecuencia, se produjo un rechazo generalizado de esta forma diabólica de transporte... Una vez que suponemos que el ferrocarril no se desarrolló –por las razones irrelevantes que fueran– podemos preguntar qué hubiera sucedido. Esto requerirá una teoría de las invenciones y el cambio tecnológico; pero ésta es la única función que cumple semejante teoría en el análisis del contrafáctico de Fogel.[33]

En respuesta a la objeción de Morriss, creo que basta reproducir un pasaje en respuesta a una objeción similar de Brian Barry[34] y desarrollada por Steven Lukes:[35]

La objeción de Lukes... desarrolla un argumento de Brian Barry y responde a mi primera objeción al argumento. Barry sostiene que podemos postular la no invención del ferrocarril aunque tengamos (y apliquemos) una teoría de que la invención era inevitable. Sostuve luego que con semejante teoría hubiéramos tenido que reinventar el ferrocarril poco después de su hipotética no invención, y Lukes dice que podemos seguir postulando su no invención: si uno puede creer en una imposibilidad, ¿por qué no en dos o en seis? La respuesta es sencilla: nuestra teoría del cambio técnico vaticinaría la invención de otro medio de locomoción, tan similar al existente como se deseara. Una teoría que permitiera vaticinar el cambio técnico implica que existe una presión en favor de una invención determinada y si uno presume la no invención pero conserva la presión, puede tener la certeza de que dará lugar a algo bastante parecido. Para reforzar el concepto, considérese otro ejemplo en el que la metáfora de la presión requiere precisión literal: ¿cómo se hubiera desarrollado la economía japonesa sin el terremoto de 1923? Es una pregunta que se puede responder si uno elimina hipotéticamente no sólo el terremoto sino también la presión que lo causó; pero si uno conservara la presión, no podría impedir que se produjera el terremoto poco después en el mismo lugar o al mismo tiempo en un lugar cercano y en tales casos los

efectos serían bastante similares. Asimismo, si uno postula la no invención tanto del ferrocarril como de sus parientes cercanos, entonces niega en realidad que exista una teoría del cambio tecnológico que permite vaticinar la aparición de sustitutos. Por eso, ratifico mi afirmación: no se puede sustentar una teoría sobre los sustitutos del ferrocarril y a la vez suponer legítimamente la no invención del ferrocarril, porque los sustitutos se parecerán tanto a la cosa verdadera que no se podrá evaluar la importancia causal de ésta.[36]

Desde luego que el cura fundamentalista de Morriss no hubiera podido postergar el ferrocarril hasta fines de siglo. En el espíritu del argumento de Lukes, podríamos invocar una sucesión ininterrumpida de obstáculos accidentales. Pero entonces tenemos el problema de identificar qué es "el ferrocarril". Supongamos racionalmente un sistema de locomoción a motor con ruedas y sobre rieles. A falta de un ferrocarril, ¿no podrían haberse desarrollado vehículos a motor que se desplazaran por caminos especiales, vedados al resto del tránsito y estabilizados por un artefacto aéreo? En términos más generales, ¿cómo se pueden admitir los sustitutos sin permitir que se acerquen arbitrariamente al sustituido? A medida que se parecen más y más, no tiene sentido hablar de una economía sin ferrocarriles. Afortunadamente, Fogel sustentó su conclusión en la suposición del peor de los casos, en el que en un mundo sin ferrocarriles no aparecieron invenciones contrafácticas. Si (como parece creer Morriss) utilizara esas invenciones para demostrar que los ferrocarriles eran prescindibles, "aserraría la rama sobre la cual está sentado", como digo en la página 254.

Notas

1. No tengo nada que agregar al capítulo 2 y en cuanto al 1 sólo cabe hacer la siguiente aclaración. G.A. Cohen señala que algunas partes de la "definición revisada" de las relaciones externas en la página 55 son redundantes, y se pueden simplificar así: *Definición revisada (revisada): Rab* expresa una relación externa si lo hace de acuerdo con la definición tentativa, y además ninguna F_i tiene una relación estructural oculta con una variable ligada, uno de cuyos valores es *b*, y ninguna *Gj* tiene una relación estructural oculta con una variable ligada, uno de cuyos valores es *a*.

2. Esto recuerda el chiste de las "dos corbatas", el mejor ejemplo del humor judío. Una madre regala a su hijo dos corbatas. Al día siguiente, cuando el joven baja a desayunar con una de las corbatas puesta, la madre pregunta: "¿Qué pasa, no te gusta la otra?" Este pertenece a la misma familia de preguntas tales como, "¿así que ya no le pegas a tu esposa?", analizadas en el capítulo 4.

3. Marx, K.: *Mathematische Manuskripte*. Kronberg, Scriptor, 1974, pág. 50.

4. Véanse *Will, Freedom and Power*. Oxford, Blackwell, 1975; y "Human abilities and dynamic modalities", en Manninen, J. y Tuomela, R., comps.: *Essays on Explanation and Understanding*. Dordrecht, Reidel, 1976.

5. Mejor dicho, una de sus dos objeciones principales. La otra es que si un mundo accesible (posible desde el punto de vista semántico) es aquel en que puedo desplegar toda

mi capacidad, semejante mundo no puede existir, porque me permitiría hacer cosas contradictorias: por ejemplo, sería capaz de fumar, capaz de abstenerme de fumar. Pero ésta es una idea muy extraña. Es mucho más concebible un mundo posible en el que puedo desplegar algunas de mis capacidades. Más extraño aún es que, según Kenny, la capacidad debería expresarse como necesidad en lugar de posibilidad, debido al uso del cuantificador universal en lugar del existencial en la definición de un mundo posible. Pero esto confunde el "todo" de la necesidad (la verdad en *todos* los mundos posibles) con el "todo" de su definición del mundo posible (un mundo en el que despliego *todas* mis capacidades). La capacidad de presentar la capacidad (y el poder) como necesidad en lugar de posibilidad es digna de estudio, pero no por esta razón.

6. *Will, Freedom and Power*, pág. 140.

7. En *The Cement of Society*, Cambridge University Press, 1989, doy varios ejemplos de la importancia de la ambigüedad deliberada en la discusión de los convenios colectivos de trabajo. [Versión castellana: *El cemento de la sociedad*. Barcelona, Gedisa, 1991.]

8. "The Constitution of the French Republic", en Marx, K. y Engels, F.: *Collected Works*. Londres, Lawrence and Wishart, 1978, vol. 10, pág. 578.

9. En otra parte (*Das Volk*, 30 de julio de 1859), Marx cita a Clausewitz para afirmar que un general que piensa demasiado en lo que hará después de lograr la victoria tendrá menores posibilidades de alcanzarla.

10. Asimismo, como se observa en el capítulo 4 y en los comentarios, el modelo epistémico de la creencia de Hintikka presupone que los creyentes son racionales.

11. Peter van Inwagen (comunicación personal) me informa sobre un modelo desarrollado por su colega Mark Brown, de la Syracuse University, en el que no se distribuye la posibilidad sobre la disyunción. Una inspección breve y evidentemente superficial del modelo indica que esto no parece tener buenos correlatos intuitivos en términos políticos.

12. Capítulo 2 de mi *Sour Grapes*, Cambridge University Press, 1983.

13. Aquí se incluye la condición de la publicidad para que los ejemplos resulten más inteligibles. Tratándose de uno mismo, es difícil no saber que cuando uno abre un libro con la esperanza de que lo haga dormir, lo hace para dormir. Por consiguiente, es difícil que uno se duerma. En cambio, un terapeuta conductista podría decirle a un paciente que sufre insomnio que trate de leer de noche y de esta manera hacerlo dormir. En el capítulo 6 de *Sour Grapes* demuestro que, en otros casos, los obstáculos a la inducción intencional de esos estados permanecen aun cuando no se afirma la intención de viva voz.

14. Quattrone, G. y Tversky, A.: "Self-deception and the voters's illusion", en Elster, J., comp.: *The Multiple Self*. Cambridge University Press, 1986, pág. 48.

15. Kesselman, J.R.: "Work relief programs in the great depression", en Palmer, J.L., comp.: *Creating Jobs: Public Employment Programs and Wage Subsidies*. Washington, The Brookings Institution, 1978, pág. 215.

16. Después de la primera edición de este libro, descubrí un ejemplo encantador que revela la necesidad de aplicar un criterio más complejo. Se dice que en una época Niels Bohr había clavado una herradura sobre su puerta. Cuando le preguntaron si de veras creía que las herraduras traían suerte, respondió: "No, pero me han dicho que traen suerte incluso a quienes no lo creen." (Segrè, E.: *From X-Rays to Quarks*. San Francisco, Freeman, 1980, pág. 171.) Si la armamos un poco, la historia es la siguiente. (1) Niels Bohr cree que "la herradura no me traerá suerte". (2) Niels Bohr cree que "las herraduras traen suerte a los que no creen que les traen suerte". No hay contradicción entre las creencias encomilladas en (1) y (2), pero si agregamos (1) en sí, percibimos intuitivamente la sensación de inconsistencia.

17. Expuse esta distinción en "Négation active et négation passive", *Archives Européennes de Sociologie 21* (1980), págs. 329-349 y, espero que en forma más satisfactoria, en el capítulo 2 de *Essais de Psychologie Politique*, de próxima aparición en Editions de Minuit, París. Allí hay mayores referencias a las obras de Zinoviev.

18. Inhelder, B. y Piaget, J.: *La genèse des structures logiques élémentaires*. París, Presses Universitaires de France, 1959, cap. V.4. [Versión castellana: *La génesis de las estructuras lógicas elementales*. Buenos Aires, Paidós.]

19. "Manndatteren og kjerringdatteren", en Asbjørnsen, P.C. y Moe, J.: *Samlede eventyr.* Oslo, Gyldendal, 1957, tomo II.

20. Veyne, P.: *Le pain et le cirque.* París, Editions du Seuil, 1976, págs. 114 y 773.

21. "On denoting", *Mind* 14 (1905), págs. 479-493.

22. Sobre la función de las presuposiciones pragmáticas en la lógica, véase Haack, S.: *The Philosophy of Logics.* Cambridge University Press, 1978.

23. Véase mi *Leibnitz et la formation de l'esprit capitaliste.* París, Aubier Montaigne, 1975, págs. 224 y sigs.

24. Cohen, G.A.: "The structure of proletarian unfreedom", *Philosophy and Public Affairs* 12, 1983, págs. 3-33.

25. Nozick, R.: "Coercion", en Morgenbesser, S. y cols., comps.: *Philosophy, Science and Method: Essays in Honor of Ernest Nagel.* Nueva York, St. Martin's Press, 1969, págs. 440-472.

26. *Democracy in America.* Nueva York, Anchor Books, 1969, págs. 651-652.

27. Gilbert, H. y Larson, E.B.: "Dealing with limited resources: The Oregon decision to curtail funding for organ transplantation", *The New England Journal of Medicine,* 21 de julio de 1988, págs. 171-173.

28. Estos párrafos están adaptados de la introducción a mi *El Cemento de la Sociedad. op. cit.*

29. El ejemplo proviene de Bicchieri, C.: "Self-refuting theories of strategic interaction" (inédito, 1987).

30. Los pasajes siguientes están basados en Binmore, K.: "Modeling rational players", *Economics and Philosophy 3* (1987), págs. 179-214, que resume la obra de Selten, Kreps y Wilson, y otros.

31. El argumento pertenece a Bicchieri.

32. Morriss, P.: *Power: A Philosophical Analysis.* Manchester University Press, 1987, Apéndice 3: "Jon Elster on Counterfactuals".

33. *Ibíd,* págs. 217-218.

34. En su reseña de esta obra en *Political Studies* 18 (1980).

35. "Elster on Conterfactuals", *Inquiry* 23 (1980), págs. 145-156.

36. "Reply to comments", *Inquiry* 23 (1980), págs. 213-232 y 222-223.

Introducción: Lógica y sociedad

Mi morada es la Posibilidad–
Una Casa más soleada que la Prosa–
Más abundante en Ventanales–
Superior —por sus Puertas–
Emily Dickinson

Las relaciones, los números, las probabilidades, el conflicto y la elección son los materiales que forman las sociedades. Las ciencias sociales tratan de ellos. En este libro intento sacar esos conceptos de su estructura abstracta y traducirlos al lenguaje más austero y transparente de la lógica formal. Un breve esquema de la teoría de la lógica se expondrá en el capítulo 1. Aquí no utilizaré ese lenguaje formal sino, más bien, intentaré dar una motivación intuitiva a quienes, quizá, consideren que éste es un extraño esfuerzo. En realidad, demando a la teoría de la lógica más de lo habitual. Me parece que todo el mundo estará de acuerdo en que el análisis lógico es una herramienta para la elaboración de teorías axiomáticas en las ciencias formales, naturales o sociales. Deseo avanzar un paso más allá, del nivel formal de la teorización al sustantivo.[1] Quiero argumentar que la teoría de la lógica puede ser aplicada no sólo en la formalización del conocimiento, obtenido por otros medios, sino también en la fase creativa y constructiva del trabajo científico. Para mencionar sólo un ejemplo, que se indaga en los capítulos 4 y 5, los epistemólogos argumentarían que la *contradicción* tiene su lugar en el metanivel, es decir, en el análisis de la forma lógica de las teorías científicas. Pienso que es posible hablar sin ambigüedades acerca de contradicciones *reales*. Por ejemplo, los fenómenos mentales o sociales –en un sentido que se especifica más adelante– pueden ligarse a la noción lógica de contradicción.

Mi argumento general para considerar el lenguaje de la teoría de la lógica como un vehículo para la investigación científica es el siguiente. La ciencia en general procede por medio de la abstracción, *desechando información*, para enfocar algún rasgo general del objeto estudiado. Para

explicar al hombre económico y su conducta en el mercado no necesitamos saber (o así lo presumimos) su color de pelo o lugar de nacimiento. Los modelos probabilísticos de la movilidad social abstraen la mayoría de los rasgos del hombre concreto, excepto la pertenencia presente o quizá pasada a una clase. Los modelos lógicos utilizados en este libro llevan ese proceso de abstracción hasta lo que creo debe ser su límite máximo. En vez de un número de individuos que varíe con el conjunto de números naturales, la lógica maneja sólo tres grados: ninguno, alguno y todos. En lugar de cuantificar las probabilidades en un continuo de 0 a 1, la lógica se ocupa sólo de posibilidades, imposibilidades y necesidades. Estas dos abstracciones son respectivamente la raíz de la lógica cuantificada y de la lógica modal. Cuando se llevan a cabo en forma simultánea nos dan la estructura de la *lógica modal cuantificada*, que es el marco de referencia básico de este libro. Tan importante como la estructura formal de la lógica modal cuantificada es la interpretación del concepto de posibilidad. La característica central de las interpretaciones más recientes es que los estados posibles son vistos no como posibles *tout court*, sino como posibles en relación con un estado dado (real). La afinidad con las ciencias sociales en este punto casi no necesita enfatizarse.

Espero, por lo tanto, que este proceso de abstracción se pueda revertir; que la percepción más aguda de la lógica abstracta de las situaciones sociales pueda llevar al científico social a formularse nuevas preguntas en un nivel más concreto. Un ejemplo típico para demostrar que la teoría de la lógica puede por lo menos sugerir nuevos conceptos es el siguiente. En el estudio de la movilidad social se describe al individuo de una clase dada como aquel que tiene varias probabilidades de transición hacia diversas clases, incluyendo la suya propia. En términos lógicos, las clases a las que él tiene probabilidades no nulas de acceder son relativamente *posibles* para la clase inicial. Las características comunes a todas las posibles clases, si las hay, podrían denominarse características necesarias del individuo en su situación inicial. Podría resultar, o no, que estas características necesarias sirvieran como variables explicativas donde las características meramente fácticas del individuo no lo hagan. Un ejemplo más conocido es que las propiedades abstractas de las relaciones –tales como simetría, transitividad y reflexividad– son frecuentemente útiles en la caracterización de las relaciones sociales, por ejemplo, el poder.

Este libro está dirigido primordialmente al científico social y sólo secundariamente al lógico profesional. Este no encontrará aquí muchas cosas de interés teórico, excepto, quizás, algunas de las ideas del capítulo 6. Sin embargo, espero que parte de este material resulte útil para fines educativos y pedagógicos. A veces se tiene la impresión de que los lógicos buscan activamente ejemplos triviales o caprichosos. Creo que la enseñanza de la lógica se beneficiaría a partir de aplicaciones más complejas y realistas, así como los ejemplos de la física son de gran ayuda en la enseñanza de la matemática. Esta comparación podría usarse para reforzar una exigencia algo mayor. Fue John von Neumann (me parece) quien dijo

que los matemáticos que no están en contacto con los problemas de la física tienden a hacerse *barrocos*, si utilizamos este término en contraposición con el estilo *clásico* de pensamiento, constantemente revitalizado por el contacto con las ciencias empíricas. En lo que alcanzo a juzgar (que no puede ser mucho, ya que no soy especialista en lógica modal), hay una cierta inclinación barroca –el embellecimiento algo inútil de las glorias pasadas– en parte del actual trabajo de los lógicos modales. Un contacto más amplio con las ciencias empíricas y, quizás, especialmente con las sociales, podría infundir nueva vida a la disciplina. Esto ya se ha dado en la lógica temporal, que ha planteado problemas (aunque no demasiado extensos) a la filosofía de la historia.[2]

Me gustaría transmitir los sentimientos hacia diversos problemas que me llevaron a escribir este libro. La fuente de inspiración más importante ha sido la prolongada controversia entre la lógica hegeliana o dialéctica y la formal. Quienes defienden la primera acusan a esta última de tratar solamente de trivialidades. Hacen una diferenciación entre la simple *comprensión* que opera en la vida cotidiana y en las ciencias y la elevada *razón* capaz de una argumentación metafísica y una percepción dialéctica. Los lógicos analíticos han respondido que la dialéctica supera la trivialidad sólo a costa de la incomprensibilidad. En los capítulos 3, 4 y 5 intento desarrollar una posición intermedia. Creo que los pensadores dialécticos han tenido un talento especial para elegir los *problemas* interesantes y a veces esenciales aunque sus esfuerzos por desarrollar un nuevo *método* sean considerados un fracaso. Tal como lo entiendo, no hay nada realmente importante en Hegel o en Marx que no pueda ser formulado en el lenguaje ordinario y en la lógica formal[3] Para que esta afirmación no sea una tautología se necesita tener una independencia de criterio acerca de lo que es realmente importante. El mío es tomar como ejemplos paradigmáticos del razonamiento dialéctico a los análisis de los filósofos dialécticos y demostrar que pueden hacerse con una argumentación lógica directa.

El lector podrá ver que la promesa del párrafo anterior sólo se refiere a *algunas* nociones dialécticas. Al final del capítulo 3 indico brevemente cómo puede interpretarse el concepto de *la negación de la negación* en los términos de la lógica modal. Sin embargo, la categoría más importante a traducir de la lógica dialéctica a la formal es la de *contradicción*. Sabemos qué es una contradicción lógica, pero ¿no es un error y un abuso del lenguaje hablar de "contradicciones reales"? A priori, pienso que habría que establecer los siguientes requisitos para formular un concepto viable de "contradicción real" en las ciencias sociales:

– Debería estar firmemente articulado con la noción de contradicción lógica,

– debería estar articulado con la teoría del cambio (individual o social),

– debería ser lo suficientemente preciso y operativo,

– debería ajustarse a algunos (por lo menos) de los escritos de Hegel y Marx acerca del tema.

La primera condición se impone por razones de claridad de expresión. "Contradicción" es un término básicamente lógico y no debería extenderse de manera que se divorcie totalmente de su significado primario. Si por "contradicción" queremos decir solamente oposición, conflicto o lucha, entonces tendremos que *decir* oposición, conflicto o lucha. Debemos resistir la tentación de abusar de la connotación lógica para hacer aparecer nuestras opiniones como interesantes, así como la de caer en las connotaciones no lógicas, de tal modo que las hagamos aparecer como posibles. El segundo requisito es razonable a la luz de la articulación que, desde Heráclito, se ha establecido entre contradicciones y cambio: "Sin contrarios no hay progresión" (William Blake). La tercera condición no necesitaría ninguna justificación, pero una pequeña observación puede servir para poner de manifiesto su intención polémica. Para que un concepto sea útil debe ser significativamente menos que totalizador. Es de poca ayuda decir, como Mao Tse-Tung, que hay contradicciones en todas las cosas. En mi definición de una contradicción real, hay una contradicción en el interior de la clase obrera pero no entre la clase obrera y la capitalista. Esta restricción, lejos de ser desventajosa, creo que ayuda a clarificar la siguiente idea. En la tradición marxista, desde Engels en adelante, ha habido una tendencia a utilizar el término "contradicción" de manera tan elástica que le quita al concepto su agudeza analítica. Valdría la pena, entonces, si fuera posible, *donner un sense plus pur aux mots de la tribu*. La última condición es la menos importante. Su cumplimiento debería ser visto como suplementario más que como un desiderátum.

No es de ninguna manera evidente que estos requerimientos puedan cumplirse simultáneamente. Podría preguntarse si la intención de hacerlo no constituye en sí misma una contradicción, en el sentido del capítulo 4. Me parece, sin embargo, que pueden ser observados de una forma que resista las críticas corrientes de los filósofos de la ciencia positivistas y analíticos. La noción general de una "contradicción real" se subdivide en dos especies: las contradicciones individuales y las contradicciones sociales (las cuales se tratan en los capítulos 4 y 5, respectivamente). La noción de una contradicción individual o mental es realmente muy sencilla, ya que se basa solamente en la idea de que una persona puede perfectamente tener opiniones o deseos contradictorios aunque éstos no puedan ser ni tornarse verdaderos. La noción de contradicción social es algo más compleja. Está ligada a la muy conocida *falacia de composición*, que yo interpreto (de manera algo diferente a como lo hacen usualmente los textos de lógica) en términos de la lógica modal cuantificada. Esta falacia puede ser utilizada como la clave para entender las dos clases de fenómenos sociales, que denomino *suboptimalidad y contrafinalidad*: el espacio entre la intención y el resultado, entre lo posible y lo real.

Intento demostrar que en verdad ambas variedades de contradicción –individuales y sociales– conducen al cambio. La articulación es brevemen-

te descrita para las contradicciones individuales y un poco más elaboradamente para las contradicciones sociales. En el capítulo 5 el lector observará sin duda que la discusión sociológica sustancial tiende a asumir tanta importancia como el análisis lógico. No pido excusas por ello, aunque saque a algunos lectores del camino. Pienso que es importante ver los conceptos *en acción* si uno desea convencerse de su fertilidad. La mera afirmación no llevará al lector a ponerse de acuerdo con la idea de que una contradicción real puede y debe ser el meollo de una teoría del cambio social. Por otra parte, no ha sido posible, por falta de espacio, presentar la teoría detalladamente, tarea que espero emprender en algún otro lugar.

Una segunda y casi tan importante fuente de inspiración de las ideas presentadas en este libro ha sido la cuestión de las afirmaciones contrafácticas en la historia. Especialmente en el capítulo 6 trato extensamente este problema rastreando la historia del concepto, discutiendo la conexión con sus actuales análisis lógicos y comentando detalladamente algunos ejemplos de construcciones contrafácticas en trabajos históricos recientes. Está suficientemente claro que las afirmaciones contrafácticas pueden ser tratadas en el contexto de la lógica modal, pero no hay un acuerdo entre los lógicos respecto a cómo exactamente analizarlas. Mi posición es que en los escritos históricos (y posiblemente en general) las afirmaciones contrafácticas deben ser entendidas en el contexto de una teoría, implícita o explícita. Esto a menudo se denomina enfoque *metalingüístico* de los contrafácticos, en contraposición al enfoque *ontológico*, que ha estado más en boga recientemente. Algunas palabras para aclarar esta distinción pueden ser procedentes, incluso en esta etapa introductoria.

Se solía decir que la historia es un tejido sin costuras. Todo lo que ocurrió tenía que ocurrir; no hay junturas donde podamos introducir una cuña y cambiar el curso de los acontecimientos, ni siquiera como un *Gedankenexperiment*. Esto puede ser cierto a nivel del determinismo universal (sin tener en cuenta las cuestiones planteadas por la mecánica cuántica), pero la ciencia –incluyendo la ciencia de la historia– no se satisface con la afirmación metafísica. Queremos saber no sólo *que* un hecho dado tiene una causa sino también *cuál* es la causa. Si no podemos especificar la causa es como si el hecho fuera incausado, que nosotros sepamos. Esto también significa que somos libres de llevar a cabo un experimento imaginario, suponiendo que el hecho en cuestión nunca ocurrió, y preguntarnos cuál sería el posterior curso de la historia. El conocimiento de las causas debe estar relacionado con la teoría; la ignorancia de la causa refleja la ausencia de una teoría. Esto implica que si en una etapa posterior del desarrollo del conocimiento sociológico se encuentra una teoría que provea la causa del hecho en cuestión, el experimento contrafáctico puede dejar de parecer legítimo. Con cada nueva teoría buena parte de la historia contrafáctica debe ser reescrita o desechada, pero también debe serlo la historia fáctica. Debe mencionarse una importante consecuencia de este punto de vista metalingüístico: el problema de

las"tijeras", creado por el doble uso de la teoría en el análisis contrafáctico. La referencia a las teorías causales es fundamental, tanto para delimitar los interrogantes contrafácticos que pueden legítimamente plantearse como para fijar nuestra capacidad de contestarlos con un grado satisfactorio de precisión. Es claro que aquí hay que lograr un equilibrio delicado: una teoría sólida elimina como sinsentidos muchas respuestas posibles, mientras que una teoría más débil puede no capacitarnos para contestar las preguntas pertinentes.

Los ejemplos presentados en el capítulo 6 son, me parece, de interés sustancial por sí mismos. La controversia sobre el nivel de vida durante la Revolución Industrial en Inglaterra es, a menudo, caracterizada en términos de la visión "optimista" versus la "pesimista", pero en realidad esto es una confusión de dos temas diferentes. Está el tema en su mayor parte (pero no totalmente) fáctico del pesimismo versus el no pesimismo: ¿se volvieron las cosas peor o no? Está, también, la cuestión contrafáctica del optimismo versus el no optimismo: ¿podrían las cosas haber sido mejores? Este último problema tiene el rasgo distintivo de ser una búsqueda del *contrafáctico óptimo*, el mejor de todos los posibles pasados. En otros ejemplos históricos el curso real del desarrollo no se contrasta con todos los demás cursos posibles sino, más bien, con una secuencia alternativa de hechos: la economía de los Estados Unidos sin el ferrocarril; la economía de la preguerra sureña sin la esclavitud. Extraigo estos ejemplos del trabajo de Robert Fogel (en el segundo caso, juntamente con Stanley Engerman); si enfoco mi discusión fundamentalmente sobre los que veo como defectos de sus análisis, esto no debe opacar el efecto inmensamente estimulante de su trabajo.

Entre los pioneros de la metodología de la historia contrafáctica menciono a Max Weber, quien vinculó nuestro interés en los pasados alternativos con nuestra libertad de elección de futuros alternativos. Si pensamos que hay una opción real *ex ante*, sería legítimo explorar *ex post* lo que hubiera ocurrido en el caso de haberse elegido cualquier otra alternativa. (No deduzco que ésta sea la única forma en que pueden justificarse los ejercicios contrafácticos.) En el capítulo 3 exploro la lógica general de tales posibilidades antes del hecho, tomando las posibilidades políticas como principales ejemplos. Este análisis puede estar intuitivamente motivado desde diversos ángulos. Primero, tenemos alguna noción intuitiva de que no todas las vías abstractas o incluso técnicas de acción son posibles *hic et nunc* en relación con un estado dado. Tomemos el siguiente comentario de Steven Lukes (*Sunday Times*, 9 de noviembre 1975) sobre lo que entonces era la situación política en Portugal: "En estas condiciones, un número de opciones debe probablemente ser descartado...Tampoco es posible en Portugal una democracia popular como las de Europa del Este, cuyo espectro ha amenazado al país hasta hace poco...Los posibles escenarios son aparentemente pocos". Estos giros del lenguaje sugieren que hay un conjunto de salidas lógicas posibles, algunas de las cuales se juzgan como políticamente posibles. Es un paso natural, entonces, preguntarse algunas cuestiones más en relación con este subconjunto políticamente posible:

¿Cuándo es irreversible el cambio político? ¿Cuándo es posible saltearse un paso en un desarrollo político? ¿Es el *status quo* siempre una opción posible? O, en términos formales, ¿la relación "políticamente posible en relación con" es simétrica, transitiva o reflexiva? Argumentaré, también, que dichos ejemplos nos permiten explorar por completo las potencialidades de la lógica modal *cuantificada* para el análisis sociológico o histórico.

Segundo, la idea de la posibilidad política es interesante desde el punto de vista formal; punto de vista, para ser precisos, que pertenece a la semántica de la lógica modal. Como se menciona antes, la revolucionaria ruptura en la lógica modal se dio con la relativización de la noción de posibilidad, lo que permitió a los lógicos dar un sentido a mucho de lo que estaba oscuro en los sistemas axiomáticos de la lógica modal. Sin embargo, ocurrió que cuando se tendría que haber ejemplificado la idea de la posibilidad relativa (con frecuencia llamada relación de *accesibilidad*), las interpretaciones resultaron triviales o no estrictamente modales. Algunos intentos fueron triviales, en el sentido de que llegaron a reintroducir el concepto absoluto (no relativizado) de posibilidad, de tal manera que se perdió todo el terreno ganado. Otros intentos fueron muy interesantes y no triviales, como la interpretación en términos de la lógica deóntica, epistémica o temporal, pero parecieron estar más bien lejanamente relacionados con la noción sustantiva de posibilidad. Por lo tanto, la interpretación política de la accesibilidad tiene la característica formal de ofrecer una interpretación que es tanto no trivial como estrictamente modal. Este no es un aporte teórico, pero la idea puede tener un valor heurístico para el profesor de lógica a quien los estudiantes le preguntan cómo puede ser que la revolución en la *lógica* modal no tenga implicación alguna en la lógica *modal*.

Tercero, me parece que la noción de posibilidad política puede ser estudiada como un caso limitante de la noción de *poder* político. El poder es comúnmente conceptualizado en términos de la probabilidad de que un actor se salga con la suya o realice una acción deseada; al ir de las probabilidades a la posibilidad pura obtenemos la idea básica del capítulo 3. También resulta que los análisis de poder y de posibilidad política se encuentran con algunos problemas iguales, especialmente la dificultad de implicar lo causal a partir del aspecto intencional. El poder no es la producción de los resultados deseados ni de los pretendidos, sino la producción inteligente de los resultados pretendidos. De manera similar, la posibilidad política se circunscribe aquí a destacar los aspectos intencional y racional de la política.

Para muchos lectores, un matiz sorprendente y sin duda enigmático será el énfasis constantemente puesto en los mundos posibles. Para el científico social puede ser siempre difícil superar la idea de que su objeto de estudio es el mundo real, mientras que las posibilidades pueden quedar, seguramente, para el novelista. Es posible dar dos respuestas distintas a esta disyuntiva, indicando dos avenidas que conducen al reino de lo posible. En primer lugar, los mundos posibles son un aspecto oculto e implícito de

toda construcción de modelos y de toda teorización. Una teoría que cubre el mundo real y sólo el mundo real no es una teoría sino una descripción. Una teoría debe implicarse en los mundos posibles, especificando el conjunto articulado de los valores realizables de las variables pertinentes. Esto se detalla en el capítulo 6. En segundo lugar, podemos referirnos a los mundos posibles para caracterizar al mundo real. Una analogía puede resaltar la idea que está detrás de esa afirmación. En los sistemas dinámicos, el ideal explicativo es predecir los valores de todas las variables en un tiempo que es $t+1$ (o la medida del cambio del tiempo t) dados solamente los valores del tiempo t. Sin embargo, en algunos casos ocurre que el análisis se desarrolla más fácilmente si admitimos los valores anteriores de las variables.[4] Basándonos en la metafísica general sabemos que el pasado no puede tener realmente una influencia causal sobre el presente, más allá de la influencia que está mediada por las huellas dejadas por el pasado en el presente. Sin embargo, puede ser conveniente (incluso, en pocos casos, necesario) hablar *como si* los valores pasados de las variables ejercieran un efecto independiente. De manera similar, puede ser conveniente admitir valores posibles de algunas variables para predecir desarrollos futuros del sistema considerado, incluso si sabemos bien que éstos podrían, *por principio*, ser reemplazados por otras variables que tomaran sólo los valores reales.

Las últimas observaciones podrían indicar que no defiendo una concepción aristotélica de las potencialidades. Las posibilidades no son entidades oscuras que rondan entre la existencia y la no existencia, y ejercen alguna clase de influencia causal sobre lo real, más allá de las entidades infinitesimales que, de alguna manera, logran ser diferentes de cero, incluso más pequeñas que cualquier magnitud que podamos nombrar. Los mundos posibles, al igual que los infinitesimales, son sólo una *façon de parler*, que pueden, en principio, ser siempre desechados. Es imposible obtener de ellos más que lo que se ha puesto en ellos. No están allá esperando una mayor inspección ni prometiendo nuevos descubrimientos. Este punto será destacado en los capítulos 3 y 6.

Mientras que los capítulos 3 y 6 intentan aportar una nueva base, el capítulo 2 es para el científico social esencialmente un recordatorio de que, al contrario de M. Jourdain, ha estado todo el tiempo haciendo poesía sin saberlo. Intento demostrar que en muchos de los muy bien conocidos análisis de las ciencias sociales tiene un lugar preminente el razonamiento posibilístico. Por lo tanto, el paso desde la matriz de las probabilidades transicionales hacia la llamada "matriz de adyacencia" en el estudio de la movilidad social es realmente un paso desde la probabilidad hacia la posibilidad. Un ejemplo de una noción posibilística (gramaticalidad) que *no* puede ser definida como una probabilidad no nula se extrae de la teoría lingüística, que parece ser *el* campo de la posibilidad pura. Sigue luego, una discusión acerca del pensamiento estructuralista francés en donde argumento que la obsesión de esa escuela por lo posible está vinculada con la búsqueda de estructuras *finitas*. La mente humana es finita, nos regoci-

jamos con lo finito, que podemos comprender. Si el conjunto posible de casos es infinito no lo podemos abarcar. Podemos caracterizar el conjunto factible si agregamos características, tales como convexidad o compacidad, o podemos escoger el mejor miembro de un conjunto de acuerdo con algún criterio de optimización, pero no podemos captar el detalle más fino de lo posible. Para los conjuntos posibles de cardinalidad finita podemos conocer todo lo que haya por conocer, de una vez para siempre. Me parece que son de gran interés filosófico los últimos dos casos discutidos en el capítulo 2, ejemplos destacados, a la vez, de razonamiento imaginativo en las ciencias sociales. El trabajo de Paul Samuelson y Amartya K. Sen muestra que la aparentemente simple noción del ingreso nacional real esconde un abismo de implicaciones contrafácticas, mientras que el creativo trabajo de C.C. von Weiszäcker sobre el cambio endógeno de los gustos abre la visión a la mente humana creando, prediciendo y reaccionando sobre sus propias posibilidades. Estos dos casos son también utilizados en capítulos posteriores.

Me gustaría cerrar esta introducción con algunas observaciones sobre el tipo de libro en el que se va a embarcar el lector. Si, como la mayoría, incluyéndome a mí mismo, prefiere libros que defienden una tesis o le ofrecen una información de difícil obtención, quedará decepcionado. Por otra parte, aun cuando una comunidad científica donde todos hubieran hecho trabajo empírico y ninguno filosofía de la ciencia seguramente sería preferida a otra donde hubiera ocurrido lo contrario (el lector, posiblemente, podrá pensar en comunidades reales que se aproximan a estos extremos), la mayoría de la gente estará de acuerdo en que la elección contra la filosofía de la ciencia no debería ser tan fuerte como para eliminarla totalmente. Tomemos, sin embargo, el siguiente argumento. La ciencia mediocre encuentra su lugar adecuado en un proceso acumulativo de investigación,[5] pero la filosofía de la ciencia mediocre, en el mejor de los casos, sería inofensiva y crearía más confusión que la que podría disipar. Por otra parte, si alguien es suficientemente talentoso para ser filósofo de la ciencia, tendría que ejercitar su don en la ciencia propiamente dicha. Se cuenta que el economista noruego Ragnar Frisch alertó a sus futuros biógrafos más o menos en los siguientes términos: si alguien escribiera mi biografía intelectual, tendría que ser tan bueno como para no perder el tiempo haciendo *eso*. Prefiero pensar que hay diferentes criterios de excelencia. La ciencia avanza gracias a gente que tiene una obsesión única, con una o dos ideas que defiende frente al sentido común y al escepticismo organizado de la profesión. Otros son mejores para producir buenas ideas que para llevarlas a cabo. Su razonamiento es tentativo e hipotético más que afirmativo y categórico. Descartes y Leibniz son buenos ejemplos de ambos casos en la historia de la filosofía. La mayor parte de la gente estará de acuerdo con que Descartes fue un pensador más importante. Personalmente me inclino por Leibniz.

Notas

1. No me considero el primero en haber hecho esto. Acerca de otros intentos en este mismo sentido, véase Kanger y Kanger (1966), Pörn (1970) o Körner (1976). Estos autores, sin embargo, escriben fundamentalmente para los filósofos y carecen (desde mi pUnto de vista) de buenos ejemplos y material concreto que puedan persuadir al sociólogo de que su trabajo *establece una diferencia*.

2. Confróntese, por ejemplo, Rescher y Urquhart (1971).

3. Esto implica un rechazo a todos los intentos de formalizar la dialéctica hegeliana en un lenguaje diferente a la teoría lógica corriente, tal como Günther (1959) o Dubarle (1970).

4. Confróntese Elster (1976a) para un más amplio análisis de algunos casos en donde la histéresis (valores pasados de las variables) ha sido integrada a las explicaciones de las ciencias sociales.

5. Véase Cole y Cole (1973, capítulo 8) para un argumento contrario.

1

La lógica formal
Una exposición informal

A todo lo largo de este libro se utiliza el lenguaje de la lógica modal cuantificada. Para el lector que no está habituado a tratar esos conceptos, este capítulo ofrece, con un mínimo aparato técnico, los prerrequisitos necesarios.[1] Los objetivos de esta exposición son, primero, familiarizar al lector con las nociones lógicas y sus interrelaciones y, segundo, bosquejar algunas de sus interpretaciones. Dejo completamente de lado todas las cuestiones atinentes a la axiomatización de los sistemas lógicos, tales como la probabilidad, decidibilidad, completud, etcétera. Estoy interesado en hacer una analogía con la lingüística, pero sólo con su estructura superficial, pues las normas generativas que constituyen la estructura profunda difícilmente son pertinentes para los objetivos de este libro. La lógica es un lenguaje y, como cualquier otro lenguaje, puede ser aprendido para su utilización o estudio. Este es un intento de dar una herramienta nueva al no lógico, y no un manual de cómo fabricar nuevas herramientas, lo que no excluye la posibilidad de llegar a un punto en el cual una mayor habilidad para la utilización de herramientas presuponga una comprensión de cómo se elaboran.

Tres conjuntos separados de conceptos forman parte de la lógica formal, que pertenecen, respectivamente, al cálculo proposicional, al cálculo cuantificacional y a la lógica de las modalidades. El segundo se construye sobre el primero, mientras la tercera puede construirse sobre el segundo o, directamente, el primero. Me ocuparé fundamentalmente de la estructura triestratificada que comprende la *lógica modal cuantificada*. Esa estructura puede ser estudiada desde dos distintos puntos de vista, cuya interrelación es el problema central de la investigación metalógica: el de la teoría *prueba* y el de la teoría *modelo* (o semántica). En la teoría *prueba* investigamos el conjunto de fórmulas que pueden ser deducidas (dadas ciertas reglas de inferencia) de los axiomas del sistema. En la teoría *modelo* nos ocupamos del conjunto de fórmulas que resultan verdaderas bajo todas las interpretaciones posibles. (Más adelante se explica qué significa aquí "interpretación". Como ya se ha afirmado, el aspecto de la teoría *prueba* de

la lógica no tiene cabida en el presente estudio, mientras que el aspecto de la teoría *modelo* es importante.

En la *lógica proposicional* tratamos con relaciones entre proposiciones de tipo atómico, cuyas estructuras internas son irrelevantes. Expresamos el carácter de tipo atómico de las proposiciones indicándolas por medio de una letra "*p*","*q*", "*r*", etcétera. Así que "*p*"es un *comodín* para una proposición como "llueve" mientras " '*p*' "es el nombre para "*p*". Utilizaré comodines indiscriminadamente, como si se trataran de una cosa real, pero tratando de no confundir las proposiciones con sus nombres. A partir de los átomos podemos construir frases moleculares, utilizando el *operador de negación y* las *conectivas "y", "o", "si entonces", "si y sólo si"*. Decimos que las frases moleculares son *funciones de verdad* de sus constituyentes, significando que el valor de verdad del compuesto depende sólo de (i), el valor de verdad de los componentes, y de (ii), la forma en que los componentes se combinan en la proposición molecular. Por ejemplo, supongamos que tenemos una proposición molecular "S_1" uno de cuyos componentes es la proposición atómica "*p*". Reemplazando "*p*" por "*q*" en todos los lugares (y suponiendo que "*q*" no está ya en "S_1"), obtenemos una nueva proposición molecular, "S_2". La exigencia de que la frase molecular sea una función de verdad de sus componentes implica que si "*p*" y "*q*" tienen un valor de verdad, también lo tendrán "S_1" y "S_2". En la lógica proposicional no queremos saber si una afirmación dada es acerca de Richard Nixon o del CO_2, sino sólo si es verdadera o falsa. Como ejemplo de una proposición compleja que *no* es una función de verdad de sus partes componentes podemos tomar "Creo que llueve". Podemos suponer que esta proposición ("S_1") es verdadera y que el componente "llueve" ("*p*") es también verdadero. Sin embargo, no se deduce que la proposición "creo que no hay un número primo mayor" ("S_2") es verdadera, incluso si el valor de verdad de la parte componente "no hay un número primo mayor" ("*q*") es igual al valor de verdad de "llueve". Un problema filosófico importante es si tales complejos no veritativos funcionales pueden ser reducidos o traducidos a complejos veritativos funcionales equivalentes (con diferentes partes componentes).

El operador de negación y las conectivas se definen como funciones veritativas, cuando dan, para todas las combinaciones posibles de valores de verdad de todos sus constituyentes, el valor de verdad del compuesto. En síntesis, esto se hace de la siguiente manera.

Negación: "$\bar{p}$" ("no *p*") es verdadero si y sólo si "*p*" es falso.

Conjunción: "*p*&*q*" ("*p* y *q*") es verdadero si y sólo si "*p*" y "*q*" son ambos verdaderos.

Disyunción: "*p* v *q*" ("*p* o *q*") es verdadero si y sólo si "*p*" o "*q*" o ambos son verdaderos.

Condicional: "$p \supset q$" ("si *p*, entonces *q*") es verdadero si y sólo si "*p*" es falso o "*q*" es verdadero.

Bicondicional: "$p \equiv q$" es verdad si y sólo si "*p*" y "*q*" tienen ambos el mismo valor de verdad.

La única definición que requiere algún breve comentario es el condicional "si *p* entonces *q* ". La definición implica que si el *antecedente* "*p*" es falso, entonces el condicional en su totalidad es siempre verdadero, sin importar el valor de verdad del *consecuente* "*q*". Para muchos, esto parece contrario a nuestras nociones intuitivas de la verdad, ya que debemos admitir la verdad de todos los condicionales con antecedentes falsos, inclusive absurdos. Tengo la vaga impresión de que esto puede ser aceptable en casos como "Si todos los humanos son racionales y de cuatro patas, entonces todos los humanos son racionales", porque hay una articulación entre el antecedente y el consecuente. La idea parece menos aceptable en los casos en que no está presente dicha articulación, como en "Si todos los humanos son cabras, entonces 2 + 2 = 5". Tendré que argumentar en un momento que la articulación lógica entre el antecedente y el consecuente es mucho mejor tratada en el primer ejemplo, bajo otro encabezamiento. Aquí sólo afirmaré que la mayor parte de los objetivos no sufrirán ningún daño si las afirmaciones del segundo ejemplo se tratan como vacías de verdad, tal como aceptamos como vacía de verdad la afirmación "Todas las A son B", incluso si no existe en la realidad ninguna cosa que sea A.

El principal caso en el que la definición del condicional sí crea problemas serios tiene que ver con los contrafácticos (también llamados "condicionales contrarios al hecho") como "Si Hitler no hubiera nacido, entonces la Segunda Guerra Mundial no habría ocurrido". Si esta afirmación es aceptable, como muy bien puede serlo, ciertamente no lo es en el sentido trivial de que está vacía de verdad. Si intentamos afirmar el valor de verdad de dichos contrafácticos asumiendo que son funciones de verdad condicionales de los valores de verdad de los constituyentes, entonces tendremos que aceptar como verdaderas, tanto la afirmación precedente como la afirmación "Si Hitler no hubiera nacido, la Segunda Guerra Mundial habría ocurrido de todas maneras", ya que ambos condicionales tienen el mismo antecedente falso. Dada esta consecuencia absurda uno puede intentar definir una función veritativa diferente para el contrafáctico o aceptar que la relación entre el antecedente y el consecuente, en este caso, no es una función de verdad. Está claro que sólo la segunda opción está realmente abierta. Para un análisis posterior, el lector debe remitirse al capítulo 6.

En este punto es útil distinguir entre *verdad y validez*. Una afirmación compleja es verdadera en relación con *una* (no necesariamente única) atribución de valor de verdad a los constituyentes, y es válida si es verdadera bajo toda posible atribución. Podemos citar como ejemplo el *dictum* de Quine "La implicación es la validez del condicional". Si "S_1" y "S_2" son dos proposiciones moleculares, decimos que "S_1" implica "S_2" si "$S_1 \supset S_2$" resultan verdaderas bajo toda posible atribución de valor de verdad a los átomos constituyentes. La razón por la cual nosotros pensamos que "Si todos los humanos son racionales y de cuatro patas, entonces todos los humanos son racionales" es aceptable, aun teniendo un antecedente falso, es que ejemplifica un condicional válido de la forma "$(p \,\&\, q) \supset q$". El caso es

que, en la mayor parte de las discusiones metodológicas, nuestro interés está en los condicionales donde hay alguna *articulación* entre el falso condicional y el consecuente, bien sea en el sentido de una articulación lógica, que hace a la validez del condicional, o en el sentido de una articulación causal, que nos hace ver la pertinencia del condicional ("Si todos los humanos fueran cabras, tendrían que comer hojas"). Sin embargo, es útil e inofensivo admitir la verdad de los condicionales con antecedentes falsos, incluso cuando dichas articulaciones están ausentes. En su definición estricta, el condicional es sólo una aseveración de la co-sostenibilidad de ciertas afirmaciones y no implica ninguna articulación formal o sustancial entre afirmaciones. Por otra parte, si el condicional es falso con un antecedente verdadero y un consecuente falso, ninguna articulación lógica o causal puede estar presente.

Una *interpretación* de una proposición compleja significa la atribución de valores de verdad a todos los constituyentes. Para una interpretación dada, una proposición compleja es verdadera o falsa. (La lógica de la verdad tiene dos valores.) Con respecto al conjunto de interpretaciones *posibles*, es válida (o tautológica), consistente o contradictoria. (La lógica de la validez tiene tres valores, una característica que comparte con la lógica modal.) Cuando más adelante tratamos sobre las *falacias*, nos referimos a los condicionales presentados con una pretensión de validez pero no realmente válidos. La no validez de un condicional no significa, sin embargo, que no puede ser verdadero a la luz de alguna interpretación, inclusive en el sentido no vacío de haber tenido un antecedente verdadero y un consecuente verdadero. Esto concluye nuestro resumen de la lógica proposicional.

En la lógica cuantificacional, dividimos el átomo para dar cuenta de su estructura interna. Estamos ahora capacitados para hablar de objetos individuales y de sus propiedades; de las relaciones entre tales objetos; del número de objetos que satisfacen algunas condiciones. Se puede hacer referencia a los individuos de dos maneras: por medio de constantes y de variables. Las *constantes individuales* "*a*","*b*","*c*"...se refieren a objetos determinados en el universo de discurso mientras que las *variables individuales* "*x*","*y*","*z*"...abarcan a todos los objetos del universo. Los predicados "*F*","*G*","*H*"... se unen bien sea a las constantes o a las variables individuales, de manera que, por ejemplo, "*Fa*" significa que el objeto *a* es *F*. Las *relaciones diádicas* "*R*","*S*","*T*"... conectan dos constantes o dos variables individuales entre sí o una constante y una variable individual entre sí, de manera que "*Rxa*" significa que "*x* tiene *relación R* con *a*". Las *relaciones N-ádicas* hacen lo mismo para un número *n* de constantes o variables. En este trabajo, todos los predicados y relaciones se toman como constantes; por ejemplo, no hay variables que abarquen el conjunto de todas las propiedades. En los sistemas lógicos de más alto nivel es posible formalizar afirmaciones como "Cada objeto tiene por lo menos una propiedad", pero en el contexto utilizado aquí, esto no puede hacerse; tampoco lo necesitaremos.

Las noción fundamental en la lógica cuantificacional son los *cuantificadores*. El *cuantificador universal* "$(\forall x)$ $(\ldots x \ldots)$" dice que es verdadero de todo x que...x..., donde los puntos deben ser llenados con algún predicado o una expresión más compleja. El *cuantificador existencial* "$(\exists x)$ $(\ldots x \ldots)$" dice que existe por lo menos una x en el universo del discurso, tal que ...x... Se dice que las variables individuales que entran dentro del alcance de un cuantificador son *ligadas* y las oraciones cuyas variables están todas ligadas son *cerradas*. En este trabajo nos ocupamos sólo de oraciones cerradas. En la teoría de la lógica, nos referimos a los individuos específicos (nombrados por constantes individuales) o a individuos en general, por medio de uno de los conceptos: "ninguno", "alguno" o "todos". Es posible expresar más finamente diferenciaciones numéricas en el lenguaje lógico, pero desde el punto de vista explicado en la introducción, la restricción a tres grados es una ventaja y no un inconveniente. Para obtener alguna estructura básica de razonamiento en las ciencias sociales, no queremos cargarnos con todo el conjunto de números naturales o reales.

Estas áridas definiciones serán trabajadas en un cierto número de ejemplos. Permítasenos primero introducir la noción de los *predicados definidos relacionalmente*, que es discutida más ampliamente para el lector interesado, en un apéndice de este capítulo. La frase "Pedro es inconquistable" se lee superficialmente como "*Fa*", pero a un nivel más profundo resulta ser como "$(\forall x)$ $(\overline{Rxa})$" en donde "*Rxa*" es la relación "x conquista a". En palabras, no hay ninguna x tal que esa x conquista a a. Al atribuirle esta propiedad a Pedro, no decimos nada acerca de su relación con alguien más en particular, pero indirectamente la proposición lo relaciona con todos los individuos en el universo del discurso. Tómese como un ejemplo más complejo el predicado relacional

$$(\forall x)\ (\exists y)\ (Rayx)\ \&\ (\forall x)\ (\exists y)\ (Ryax).$$

Al leer "*Rxyz*" como "x es mejor que y en z", vemos que esta expresión definida relacionalmente revela el predicado monádico "un sujeto que hace de todo pero nada bien"o, por lo menos, un componente importante de ese predicado.

La utilización de cuantificadores anidados (*nested*) (cuantificadores que están ubicados dentro del campo de algún otro) en el último ejemplo es un rasgo esencial de la lógica de la cuantificación.[2] Como muchas de las diferenciaciones en el presente trabajo están dirigidas al *orden* en el cual los cuantificadores anidados (y otros operadores) aparecen en la frase, se darán algunos ejemplos que familiaricen al lector con este problema. Para demostrar lo fundamental que es el orden de los cuantificadores podemos tomar como ejemplo simple la prueba clásica de la existencia de Dios. El ejemplo se basa en una confusa falacia entre

$(\forall x)$ $(\exists y)$ (Rxy)– todo tiene una causa,
$(\exists y)$ $(\forall x)$ (Rxy)– hay algo que es la causa de todo.

Donald Davidson ha argumentado que la misma confusión está presente cuando deducimos el reduccionismo del materialismo y el determinismo. "Si se aplica un determinado concepto psicológico a un hecho y no a otro, debe haber una diferencia que pueda ser descrita en términos de la física. Pero no debe concluirse que hay una sola diferencia física descriptible que distinga dos hechos cualesquiera que difieran en un aspecto psicológico dado."[3] Es decir, Davidson nos alerta que escribir "*Rxy*" por "*x* cae bajo la descripción *y*", "*Fy*" por "*y* es una descripción psicológica" y "*Gy*" por "es una descripción física", puede llevar a confundir

$$(\forall v)\ (\forall x)\ (\forall z)\ (\exists w)\ ((Rxy\ \&\ \overline{Rzy}\ \&\ Fy) \supset (Rxw\ \&\ \overline{Rzw}\ \&\ Gw))$$

con

$$(\forall y)\ (\exists w)\ (\forall x)\ (\forall z)\ ((Rxy\ \&\ \overline{Rzy}\ \&\ Fy) \supset (Rxw\ \&\ \overline{Rzw}\ \&\ Gw)).$$

Un ejemplo similar de las ciencias sociales es la diferenciación entre la división personal y funcional del trabajo, por ejemplo, entre la distribución de personas en los trabajos y el sistema de trabajo mismo. Si se permite que "*Rxyt*" represente "la persona *x* ocupa la posición *y* en el tiempo *t*", dos sociedades diferentes pueden ser descritas respectivamente por

$$(\forall x)\ (\exists y)\ (\forall t)\ (Rxyt), \qquad (\forall x)\ (\forall t)\ (\exists y)\ (Rxyt).$$

La primera sociedad es muy distinta de la segunda, aunque en un momento dado no puedan distinguirse. Sólo un enfoque longitudinal, que vea la razón de los cambios de personal en el trabajo como parte de la estructura social, puede captar esta diferencia. En *La ideología alemana*, Marx parece cometer la falacia de afirmar que una sociedad con una división funcional del trabajo debe también tener congelada para las personas (ejemplo, para toda la vida) la división del trabajo,[4] aunque en *El capital*, la diferenciación entre manufactura e industria es fundamental.[5]

Un último ejemplo es la interpretación del famoso teorema de la imposibilidad de Lincoln. Dando "*Rxt*" por "usted puede engañar a una persona *x* en un tiempo *t*", el teorema puede ser entendido de dos maneras diferentes:

$$(\forall x)\ (\exists t)\ (Rxt)\ \&\ (\exists x)\ (\forall t)\ (Rxt)\ \&\ \overline{(\forall x)\ (\forall t)\ (Rxt)},$$
$$(\exists t)\ (\forall x)\ (Rxt)\ \&\ (\exists x)\ (\forall t)\ (Rxt)\ \&\ \overline{(\forall x)\ (\forall t)\ (Rxt)},$$
$$(\forall x)\ (\exists t)\ (Rxt)\ \&\ (\forall t)\ (\exists x)\ (Rxt)\ \&\ \overline{(\forall x)\ (\forall t)\ (Rxt)},$$
$$(\exists t)\ (\forall x)\ (Rxt)\ \&\ (\forall t)\ (\exists x)\ (Rxt)\ \&\ \overline{(\forall x)\ (\forall t)\ (Rxt)}.$$

Cuando también introducimos en la lógica modal los operadores modales, las falacias resultantes de los cambios cuantificador/operador se suman a la lista de falacias. Estas falacias, especialmente la *falacia de*

composición, son un tópico central de este trabajo, en particular del capítulo 5. Es importante, sin embargo, que el lector comprenda muy bien las falacias creadas por los cambios de los cuantificadores más simples. Cuando un cuantificador universal precede al existencial, se realiza la afirmación débil de que para cada objeto de un tipo (el tipo ligado por el cuantificador universal) puede encontrarse algún objeto de otro tipo (el tipo ligado por el cuantificador existencial) que satisface determinadas relaciones. Si se revierte el orden de los cuantificadores, la afirmación se fortalece muchísimo, porque ahora afirmamos que un objeto del segundo tipo hará el trabajo para todos los objetos del primer tipo.[6]

Las nociones de interpretación y validez en la lógica cuantificacional son algo más complejas que en el caso proposicional. Primero, una interpretación requiere una especificación del universo de discurso, por ejemplo, el conjunto de objetos que constituirá el rango de las variables. En segundo lugar, debemos dar interpretaciones extensionales de los predicados y las relaciones; por ejemplo, tenemos que especificar para cada predicado y cada objeto si el predicado puede ser con verdad afirmado acerca del objeto y para cada relación *n*-ádica y para cada *n*-tupla ordenada de objetos si se da la relación entre estos objetos. La validez es entonces definida como verdad bajo todas las posibles interpretaciones, lo que significa toda especificación posible en todos los posibles universos de discurso. (En verdad, se puede demostrar que sólo el *número* de objetos en el mundo es relevante, de tal manera que una fórmula válida es la que resulta ser verdadera, bajo toda especificación extensional posible en los universos de todas las cardinalidades posibles.) Con esto concluye la explicación de la lógica cuantificacional.

En la lógica modal introducimos dos operadores adicionales, "M" y "N" cuyas interpretaciones corrientes son posibilidad y necesidad, respectivamente. Los operadores se adosan a afirmaciones completas, sean moléculas o átomos del cálculo proposicional, o totalidades estructuradas en el cálculo cuantificacional. Por lo tanto, una oración "*N*(*p*v*q*)" debe leerse "Es necesariamente verdadero que *p* o *q* es el caso" y "*M*(∃*x*) (*Rax*)", como "Es posible que exista una *x*, tal que *a* tiene *R* con *x*". (Las lecturas no estándar de los operadores serán discutidas más adelante.) La sintaxis y la semántica de la lógica modal son, sin embargo, más abstrusas que en los casos discutidos anteriormente. En el aspecto sintáctico, es difícil establecer los axiomas que gobernarán las denominadas *modalidades mixtas e iteradas*, tales como "*MNMp*" o "*NNp*". Desde 1912, se han discutido diversos postulados y uno de ellos afirma que no puede haber simplificación de las modalidades mixtas e iteradas; otro (S-4), que las modalidades iteradas deben ser reducidas a modalidades simples (*MM* siendo equivalente a *M*, *NNN* a *N*, etc.); un tercero (S-5), que *cualquier* cadena de modalidades (mixta o iterada) es equivalente al operador que más aparece (*MNNMN* siendo equivalente a *N*, etc.).[7]

El motivo por el cual no se hizo ninguna elección firme entre estos sistemas axiomáticos es que el aspecto semántico se encontraba en un

estado igualmente incierto. La única interpretación de posibilidad y necesidad bien establecida fue la que propuso Leibniz: la posibilidad es verdad en algún mundo posible, la necesidad es verdad en todo mundo posible. Estas nociones coinciden con las de consistencia y validez, antes definidas. Sin embargo, la dificultad de esta interpretación es que anula todas las modalidades mixtas e iteradas y las convierte en modalidades simples, de manera que sólo queda el sistema, más bien trivial, S-5. Quizás el problema subyacente puede ser formulado en los siguientes términos: para poder obtener estructuras formales interesantes, los lógicos querían hacer comprensibles las modalidades mixtas y/o iteradas, pero la única interpretación posible a mano tenía como consecuencia que las anulaba y convertía en modalidades simples.

La revolución en la lógica modal ocurrió a fines de la década de 1950, cuando varios autores propusieron independientemente una *relativización de la noción de mundos posibles.* En vez de decir que un mundo es posible tout court hoy decimos que es posible en relación con otro mundo. Como sinónimo de posibilidad relativa utilizamos el término más sugestivo "accesibilidad"."*Np*" se toma, entonces, para significar que "*p*" es verdadera en todos los mundos posibles en relación con el mundo real. "*NNp*" recibe una interpretación natural, al efecto de que "*p*" sea verdadera en todos los mundos posibles en relación con el real. Expresado en palabras, "*p*" es necesariamente verdadera (en nuestro mundo) si es verdadera en todos los mundos accesibles en dos pasos desde el real. De manera semejante, una modalidad combinada como "*MNp*" hoy significa que existe por lo menos un mundo *x* que es posible en relación con el real, tal como *p* es verdadera en todos los mundos posibles en relación con *x*. Los diversos sistemas axiomáticos propuestos podrían ahora ser clasificados de acuerdo con las condiciones lógicas que imponen a la relación de accesibilidad. Con una relación reflexiva, transitiva y no simétrica de posibilidad relativa, obtenemos el sistema S-4; con una relación reflexiva, transitiva y simétrica, S-5. Actualmente, de manera natural, han surgido otros muchos sistemas posibles. En verdad, podrían hacerse sistemas a medida para los más variados objetivos.[8]

Desde el punto de vista de la teoría modelo fue en verdad satisfactorio poder darle un sentido semántico distinto de S-5 a diversos sistemas y construir sistemas nuevos, anteriormente impensables. Sin embargo, la idea de posibilidad relativa sólo es un instrumento algebraico conveniente hasta que no hayamos especificado, en casos concretos, lo que dicha posibilidad relativa podría significar. Los sistemas de lógica modal son interpretados por medio de los así llamados sistemas modelo o conjuntos de conjuntos de oraciones que satisfacen diversas condiciones. Para formular algunas de estas condiciones, la noción de posibilidad relativa es fundamental, pero sólo como una relación entre conjuntos de oraciones; no está implicada una especificación concreta de posibilidad. El intento de proveer esa especificación ha resultado en la situación paradójica siguiente. O bien la especificación propuesta nos lleva de nuevo a S-5 y a Leibniz, de

manera que todas las modalidades mixtas e iteradas se anulan y perdamos todo el terreno aparentemente ganado, o bien la especificación puede metafóricamente ser llamada concepto *modal*, en el sentido clásico. Repasaré brevemente la primera interpretación y me detendré más largamente en la segunda.

La especificación más natural sería decir que un mundo es posible en relación con el nuestro si obedece a las mismas leyes lógicas o causales. En este caso, sin embargo, la utilización esencial de la palabra "mismas" en la definición hace de la relación de accesibilidad una relación de equivalencia, es decir, una relación que es reflexiva, simétrica y transitiva. El sistema resultante es, por lo tanto, S-5. No es enteramente correcto decir que la necesidad surge como verdad en todos los mundos posibles, como fue el caso en el análisis original leibnizano. Más bien una afirmación es necesariamente verdadera si y sólo si es verdadera en todos los mundos de la clase de equivalencia a la que pertenece nuestro mundo bajo la relación de equivalencia. Por lo tanto, ni la necesidad lógica ni la necesidad física son candidatas posibles para el tipo de estructura –modal y no trivial– que tenemos en mente. Son modales pero triviales.

Para candidatos a lo no trivial y lo no modal nos dirigimos a la lógica deóntica, epistémica o temporal, que son interpretaciones no corrientes de los sistemas axiomáticos de la lógica modal. La lógica epistémica es discutida más o menos detalladamente en el capítulo 4. Aquí sólo anotaremos que se da en dos versiones, una en la cual el operador "*N*" es interpretado como conocimiento, y otra en la cual es entendido como creencia. En la lógica temporal podemos interpretar "*Np*" como "es verdadero en todos los tiempos (pasado, presente y futuro) que *p*" y "*Mp*" como "es cierto en algún tiempo (pasado, presente o futuro) que *p*". Otra versión toma "*Np*" como "es y siempre será el caso que *p*" y "*Mp*" como "es o será el caso que *p*". La primera versión sólo da S-5 mientras que la segunda nos deja con S-4, porque "*MNp*"no colapsa en "*Np*". Incluso, si desde alguna fecha futura en adelante siempre será verdadero que *p*, esto no implica que en este momento *p* sea el caso. Versiones ulteriores de lógica temporal han sido trabajadas, muchas de ellas por el extinto A.N. Prior.[9]

En la lógica deóntica hay una analogía obvia (volviendo a Leibniz) entre la necesidad y la obligación, y entre la posibilidad y la permisibilidad.[10] Aquí puede verse fácilmente que ni S-4 ni S-5 tomarán nuestras nociones intuitivas y, de hecho, debe desarrollarse un sistema bastante complicado para no caer en situaciones antiintuitivas. Entre las propiedades más interesantes del sistema, podemos anotar que la relación de accesibilidad no es reflexiva. El mundo real no es necesariamente una alternativa ética para sí mismo, aunque es una alternativa para sí mismo si existe otro mundo con respecto al cual es alternativo —una propiedad a la que volveré en el capítulo 3, bajo el encabezamiento "reflexividad limitada"—. Además de los sistemas "puros" de la lógica deóntica, temporal o epistémica uno puede imaginar sistemas "combinados" para la formalización de oraciones tales

como "creo que será el caso que p" o "es obligatorio conocer que p"; mayores extensiones pueden dar "deseo que p" o "espero que p". Se ha trabajado poco acerca de estos problemas y no trataré de abordarlos aquí.

En el presente trabajo, se clarifican dos especificaciones no triviales de los operadores modales. En el capítulo 4 se hace uso extensivo de la lógica epistémica, mientras que el capítulo 3 introduce el operador de posibilidad política. El último tiene la ligeramente interesante característica de ser tanto no trivial como estrictamente modal. Decir que la relación de posibilidad política es reflexiva, simétrica y transitiva implicaría que el *status quo* es siempre una opción factible, que el cambio político es siempre reversible y que lo que puede hacerse en dos pasos puede hacerse siempre en uno, que son presunciones por cierto controvertibles. No argumento que exista un sistema correcto de posibilidad política pero pienso que la política, en el sistema S-5, es una noción muy poco interesante, que puede ser descartada sin mayor discusión, así que el sistema en realidad es no trivial.

En lo que atañe a la interpretación y la validez, la situación es mucho más compleja que en los casos precedentes. En la lógica modal no hay una noción general de validez que abarque todas las posibles interpretaciones. Más bien hay una noción de validez para cada interpretación, que se especifica indicando (i) las propiedades formales de la relación de accesibilidad, (ii) los tamaños relativos de las poblaciones en los diversos mundos posibles y (iii) la forma en la cual decidimos asignar valores de verdad a las afirmaciones y negaciones que conciernen a los objetos no existentes. En vez de entrar en detalles sobre estos problemas que muy pronto se tornan prohibitivamente complejos,[11] daré un ejemplo bien conocido de una fórmula cuya validez depende fundamentalmente de la forma en que especifiquemos los diversos aspectos arriba mencionados. (En el capítulo 3 se dan otros ejemplos.) Esta es la así llamada fórmula de Barcan, la cual puede ser escrita de dos maneras alternativas y equivalentes:

$$(\forall x)\ (N(Fx)) \supset N(\forall x)\,(Fx), \qquad M(\exists x)\,(Fx) \supset (\exists x)\,(M(Fx)).$$

Antes de mirar la interpretación de estas fórmulas, puede ser de utilidad para el lector no entrenado ver exactamente en qué sentido son equivalentes. (El lector puede saltearse el resto de este párrafo si piensa que está suficientemente familiarizado con la manipulación de los conceptos lógicos.) Primero observamos que los dos cuantificadores son definibles en términos el uno del otro, como también los dos operadores modales: $(\exists x)\ (Fx) \equiv \overline{(\forall x)\ (\overline{Fx})}$ y $Mp \equiv \overline{N\bar{p}}$. En segundo lugar, un conocido teorema (e intuitivamente obvio) del cálculo proposicional, afirma que $(p \supset q) \equiv (\bar{q} \supset \bar{p})$. Esto significa que (dando "Gx" por "$\overline{Fx}$"), los siguientes son todos equivalentes entre sí:

$$\begin{array}{l}(\forall x)\ (N(Fx)) \supset N(\forall x)\,(Fx),\\ \overline{N\,(\forall x)\ (\overline{Fx})} \supset \overline{(\forall x)\ (N(Fx))},\\ M\,\overline{(\forall x)\ (\overline{Fx})} \supset (\exists x)\ \overline{(N(Fx))},\end{array}$$

$M\,(\exists x)\,(\overline{Fx}) \supset (\exists x)\,(M(\overline{Fx}))$,
$M\,(\exists x)\,(Gx) \supset (\exists x)\,(M(Gx))$.

Por lo tanto, si la primera de estas inferencias es válida, entonces también lo es la última, y viceversa. La primera afirma que si todos los objetos en el mundo real tienen la propiedad F en todos los mundos posibles con respecto al real, entonces todos los objetos en todos los mundos posibles con relación al real tienen la propiedad *F*. Una condición suficiente para que esto sea verdadero es claramente que el mundo dado y todos los mundos que sean accesibles desde él tengan exactamente los mismos objetos. Si uno de los mundos accesibles contiene un objeto que no es parte del mundo real, la fórmula de Barcan se destruye, si ese objeto es no *F*. Otra dificultad que surge con frecuencia al interpretar las fórmulas modales es un problema planteado primero por Bertrand Russell: ¿Cómo analizaremos una afirmación como "El actual rey de Francia no es calvo"?[12] ¿Niega la existencia de un rey calvo o afirma la existencia de un rey no calvo? La primera lectura es verdadera, la segunda, falsa. Vuelvo a este problema en el capítulo 3.

La segunda (y más corriente) versión de la fórmula de Barcan se discute frecuentemente en términos de la diferenciación entre las modalidades *de re* y *de dicto*. Puedo estar muy dispuesto a admitir (*de dicto*) que hay un dronto* sobreviviente con una pata coja, pero no por ello me siento obligado a aceptar la existencia de un dronto que probablemente (*de re*) tenga una pata coja. Para hacer esta diferenciación, tendremos que rechazar la fórmula de Barcan. Esto significa que en las interpretaciones en las que queremos que la fórmula de Barcan surja como no válida es necesario tener cuidado de especificar las reglas que gobiernan la accesibilidad, los individuos y los no existentes, para obtener este resultado. Por ejemplo, en la lógica epistémica, discutida más adelante en el capítulo 4, podemos tomar la inferencia que va de "Todo lo que es un dragón se cree que es verde", a "Se cree que todos los dragones son verdes". Incluso si nos limitamos al caso de la creencia racional, la conclusión evidentemente no se sigue, ya que requeriría la premisa adicional de creer que aquellos, digamos, doce dragones son todos los que hay.[13]

Otro ejemplo de la inferencia falaz de la modalidad *de dicto* a partir de la modalidad *de re* es la falacia de composición, discutida ampliamente en los capítulos 3 y 5. Al igual que en la fórmula de Barcan esta inferencia falaz existe en dos variantes, alternativa y equivalente:

$$N\,(\exists x)\,(Fx) \supset (\exists x)\,(N(Fx)), \qquad (\forall x)\,(M(Fx)) \supset M(\forall x)\,(Fx).$$

Un contraejemplo que desaprueba el primer cambio del operador/cuantificador es el siguiente: en una república necesariamente hay alguien que es presidente, pero no hay uno que necesariamente sea presidente. Esta y otras modalidades políticas son examinadas en el capítulo 3. Como

* Dronto: ave antigua, ya desaparecida [T.]

contraejemplo de la segunda, tómese la inferencia de la posibilidad de sobreproducción simultánea en todos los sectores a partir de la posibilidad de sobreproducción en un solo sector de una economía (no monetaria). Este y muchos otros ejemplos son discutidos en el capítulo 5. De nuevo, es sencillo encontrar ejemplos de interpretaciones no corrientes. De "Me parece que usted está pensando acerca de algo" no se puede inferir que hay algún objeto específico tal que yo crea que usted está pensando en *él*. Con frecuencia puede ser obligatorio escoger *algún* curso de acción aunque la elección entre posibles vías sea moralmente arbitraria; frecuentemente se dice que en verdad ésta constituye la esencia de la decisión militar. Estos ejemplos sugieren que la inferencia en cuestión resulta invalidada en *todas* las interpretaciones de la lógica modal, y puede de veras demostrarse que éste es el caso.[14]

En lo que concierne a los individuos y sus propiedades en los mundos posibles donde existen, la discusión filosófica reciente ha planteado muchos interrogantes profundos, relacionados con antiguas cuestiones acerca del esencialismo. Algunos de estos temas son discutidos en el apéndice de este capítulo, mientras otros son una parte importante de los análisis de los capítulos 3 y 6. Aquí sólo afirmaremos, sin argumentación, la necesidad de una teoría *genética* de la individuación. Una condición, necesaria y suficiente, para que algún individuo, en un mundo posible, sea la misma persona que un individuo dado en el mundo real es que tenga los mismos ancestros en el mundo real. Esta solución al problema de la "identidad transmundo" tiene diferentes ventajas sobre la alternativa principal, que es ceder en los criterios de identidad para permitir a un individuo dado tener diversas "contrapartes" en un mundo posible dado.

APENDICE DEL CAPITULO 1

Lógica relacional y relaciones sociales

Tómense dos imágenes de la sociedad: la teoría marxista de las clases y la teoría de la estratificación social. La primera utiliza un solo criterio (económico) para la definición de las unidades sociales básicas, la segunda utiliza criterios múltiples (ingreso, educación, status, poder). La primera se empeña en explicar la distribución de atributos distintos a los de clase como una función de la diferencia de clases, tanto a nivel de la estructura como del individuo.[15] La segunda comienza con un conjunto de clasificaciones y luego avanza en la búsqueda de las correlaciones entre las diversas dimensiones. La primera teoría ve las desigualdades sociales como la causa fundamental del cambio y es, por lo tanto, conducida a ver la "inconsistencia del status" o "el desequilibrio de la posición social" como un obstáculo para el cambio. La segunda tiende a ver la desigualdad social como un hecho universal e inevitable que no tiene potencial para el cambio per se y, por lo tanto, la agitación social es explicada por la persistencia de las inconsistencias

de status, presumiendo que la gente tiende a actuar buscando igualar sus posiciones en las diversas dimensiones de la estratificación. La primera teoría sostiene una posición realista en la cuestión de la existencia de clases, mientras que para la segunda las clases existen sólo para la mirada de quien las observa. La primera teoría mira la *interacción* de las clases, la segunda, las *diferencias* de clase.

En este apéndice exploraré algunos problemas lógicos relacionados con la última proposición. Discutiré la tesis[16] de que el marxismo y las teorías "burguesas" de clase difieren en cuestiones lógicas fundamentales: la primera es una teoría de las relaciones internas de clase y la segunda postula sólo relaciones externas entre las clases. Argumentaré que la interacción o la diferencia-comparación en verdad pueden ser analizadas en términos de la lógica relacional, pero que no coinciden plenamente con la diferenciación interno-externo. Sin embargo, antes de entrar en los detalles del análisis, quisiera proponer una redefinición de las oposiciones bosquejadas en el párrafo anterior. Por un lado agrupo todas las teorías que usan conceptos interaccionales en la definición de las unidades sociales básicas, de otro lado junto todas las teorías que utilizan conceptos diferenciales. En la esfera económica ésta es una diferenciación entre la teoría de la explotación y la teoría de los grupos de ingreso. Las clases explotadoras y explotadas se relacionan entre ellas por interacción; no sólo "están relacionadas unas con otras, sino que se relacionan activamente".[17] En contraste, los grupos de ingreso están relacionados unos con otros por la relación entre mayores o menores ingresos, que es la relación de diferencia o comparación, más que de interacción. Es bien conocido que Marx rechazaba esta concepción cuantitativa de clase.[18]

En el interior de la dimensión política se puede utilizar la misma diferenciación para clasificar las diversas teorías del poder propuestas en los últimos años. De un lado hay una aproximación interaccional que ve al poder básicamente como "poder sobre alguien", concepción que implica una diferenciación cualitativa entre quienes detentan el poder y los objetos de ese poder. De otro lado hay una aproximación diferencial que atribuye a cada persona o grupo una cierta *cantidad* de poder y sobre esta base luego pasa a distinguir (cuantitativamente) entre los grupos. En el último caso, el poder no es poder por encima de alguien, del mismo modo que el dinero no es dinero por encima de alguien. La noción diferencial de poder es algo similar a adquirir el poder, una especie de ductilidad para obtener lo que se desea. La noción de interacción es análoga a la de explotación, en que implica la existencia de un individuo o grupo polares que son objetos del poder. Sostengo que en esta tipologización de las teorías de clase o estratificación, la diferenciación entre los criterios económico y político de clase es menos importante que la distinción entre los criterios unidimensional y multidimensional y también menos importante que la diferenciación entre las teorías interaccional y diferencial. Dahrendorf está más cercano a la tradición marxista que Wright Mills, aunque Wright Mills está realmente más cercano a esa tradición que Parsons.

Hay una diferenciación implícita en lo dicho arriba entre las siguientes afirmaciones relacionales:

a explota a *b*	(1)
a tiene más dinero que *b*	(2)
a tiene poder sobre *b*	(3)
a tiene más poder que *b*	(4)

Podemos también decir

a se ubica a sí mismo más arriba que *b*	(5)
a tiene más prestigio que *b*	(6)

De éstos, (1), (3) y (5) representan interacciones, mientras que (2),(4) y (6) representan comparaciones. La tarea de este apéndice es explicar esta diferencia en términos lógicos. La idea de diferenciar entre dos tipos de relación se remonta a Leibniz, quien escribió que "las relaciones son relaciones de comparación o de conexión".[19] La terminología de las relaciones internas y externas es más reciente, pero el significado de los términos está lejos de ser bien establecido.[20] Igualmente mal establecidas están las teorías que diversos filósofos han intentado desarrollar por medio de estos conceptos. Hasta donde alcanzo a ver, es posible diferenciar tres afirmaciones fundamentales que han sido expuestas: todas las relaciones son internas; todas las relaciones son externas; en la naturaleza inanimada todas las relaciones son externas, mientras que en la sociedad todas las relaciones son internas. Discutiré fundamentalmente la segunda mitad de la tercera afirmación. En mi análisis de las nociones acerca de las relaciones externas e internas, esta mitad de la afirmación resulta ser verdadera, aunque la primera parte no parece correcta.

Mencionaré, más bien para una referencia posterior, el enfoque tradicional de las relaciones externas e internas.[21] En la mayor parte de las discusiones filosóficas, la diferenciación ha sido vista como similar a la diferenciación entre las propiedades accidentales y esenciales: las relaciones internas son relaciones en las que uno o ambos relata podrían no haber sido lo que son, mientras que las relaciones externas son tales que podemos abstraer a partir de ellas en el pensamiento sin que los relata pierdan su identidad. Si queremos decir, en un terreno intuitivo, que la relación "ser más alto que" es una relación externa esto parece permitido en esta aproximación. Si, de otro lado, nuestras intuiciones lingüísticas (para lo que son meritorias[22]) nos dicen que "estar enamorado de" es una relación interna, no es evidente u obvio que esto puede volverse correcto equiparando las relaciones internas a las esenciales. El que yo esté enamorado puede ser lo que hace de mí el individuo que soy pero también puede ser lo que me hace una persona diferente:

Ahora me has amado un día entero,
Mañana cuando te vayas, ¿qué dirás?

¿Antepondrás algún compromiso nuevo?
¿O dirás que ahora
No somos las personas que éramos?[23]

Como fue señalado por G. E. Moore en una contribución clásica al problema,[24] es trivial decir que si *a* tiene una relación *R* con algo, entonces cualquier cosa que no tenga relación con este algo no puede ser *a*. El problema real es contrafáctico, si alguien que no hubiera tenido esa relación con algo no podría haber sido *a*. Sabemos que todas las personas que no tienen a John F. Kennedy como hermano no son Edward Kennedy, pero tal vez necesitaríamos saber si Edward Kenendy podría haber sido Edward Kennedy sin haber tenido a John F. Kennedy como hermano. Para contestar la última pregunta tendríamos que definir los criterios de identidad de las cosas, para que las cosas sean lo que son. Una cuestión difícil, en verdad. Podemos definir fácilmente las propiedades o relaciones necesarias para que una silla sea una silla, pero no lo que se requiere para que una silla sea *esa* silla en particular. El hecho mismo de que sea tan sencillo para nosotros poner "esa" en itálica para darle el sentido de individualidad es evidencia de la necesidad profundamente sentida que tenemos de referirnos a los particulares de una manera que sea independiente de toda propiedad o relación específica. Como argumentó Hegel en el primer capítulo de la *Fenomenología del espíritu*, cualquier objeto puede ser "ese"objeto, así que la expresión que utilizamos para la identificación de particulares es la más universal de todas las expresiones. Los particulares puros no se podrían diferenciar y serían por lo tanto idénticos.

En los capítulos 3 y 6 consideraremos la solución propuesta recientemente a este problema por Saul Kripke y Michael Dummett. Sin embargo, aquí obviaré la cuestión utilizando una aproximación diferente y puramente formal a la diferenciación entre las relaciones externas e internas, de acuerdo con la cual no hay ninguna necesidad de apelar a la pretendida distinción entre relaciones esenciales y accidentales. Si alguien pensara que con esto hago una ruptura arbitraria con la tradición filosófica sólo puedo contestar (i) que la definición aquí propuesta tiene la ventaja de ser tolerablemente precisa y (ii) que por lo menos parece adecuada a algunas de nuestras intuiciones lingüísticas (para lo que son meritorias, una vez más). Mi punto de partida será algún trabajo filosófico reciente sobre la teoría de las relaciones en la filosofía de Leibniz.[25]

Un problema fundamental en la interpretación de Leibniz es que él parece sostener dos teorías diferentes e incompatibles acerca de las relaciones entre las mónadas. Por un lado, parece argumentar que todas las afirmaciones relacionales pueden ser reducidas o inferidas de afirmaciones no relacionales[26] pero, por otro lado, afirma bastante explícitamente que todos los predicados monádicos en un análisis ulterior resultan tener una estructura relacional.[27] La solución que hace posible reconciliar estas dos tesis (propuesta por Hintikka) nos ayudará también en el análisis de las

relaciones externas e internas en las ciencias sociales. Comencemos con una definición tentativa de relaciones externas:

Definición tentativa: *Rab* es una relación externa si existen predicados (no necesariamente únicos) $F_1...F_n$ y $G_1...G_m$ tales que "*Rab*" pueda ser inferida de una función veritativa de "F_1a", "F_2a",..."F_na", y "G_1b", "G_2b"..."G_mb".

En esta definición "*a* es más rico que *b*" expresa una relación externa porque está implicada por "*a* tiene 10 dólares y *b* tiene 5 dólares". Por otro lado,"la partícula *a* es atraída por la partícula *b*" expresa una relación interna, una conclusión que no entra en violento conflicto con nuestras intuiciones. Estos dos ejemplos demuestran evidentemente que la adjudicación de las relaciones internas sólo a las ciencias sociales y de las relaciones externas sólo a las ciencias naturales no puede ser correcta. La cuestión, sin embargo, no es así de simple. Leibniz argumentaría que en el segundo ejemplo puede probar mediante un razonamiento metafísico general que la relación de atracción puede ser reducida de la forma indicada por la definición. Para el objetivo presente, sin embargo, creo que podemos obviar este problema y sólo decir que mientras la reducción no haya sido *efectivamente* lograda, la relación de atracción debe ser considerada como interna. Pueden hacerse objeciones –de naturaleza más seria– a la interpretación de "*a* es más rico que *b*" como una relación externa. Considérense, por ejemplo, las afirmaciones siguientes:

a tiene más poder que *b*, (4)
a es más alto que *b*. (7)

De acuerdo con la definición tentativa, ambas son relaciones externas, pero los predicados reductores *F* y *G* son bien diferentes en ambos casos. El predicado "*n* centímetros alto" es un predicado reductor que es en sí mismo irreductible, es decir, no contiene elementos relacionales ocultos. Pero "tener *n* grados de poder" *es* un predicado relacional implícito, que no puede ser definido sin referirse a otras personas u objetos, o a ambos. Esto puede hacerse en formas diferentes en diferentes modelos de poder,[28] pero un concepto puramente monádico de poder seguramente sería un absurdo. Si volvemos a Leibniz, podemos interpretarlo diciendo que sostenía la siguiente tesis: incluso si una afirmación relacional "*Rab*" es reducida de manera que pueda ser inferida de, digamos, "*Fa* & *Gb*", esto no excluye que los predicados reductores "*F*" y "*G*" puedan tener en sí mismos una estructura relacional, tal como "ser el amo de todos" o "ser el esclavo de nadie".[29] En efecto, Leibniz sostenía una tesis más fuerte: que *todas* las afirmaciones "*Rab*" pueden ser reducidas a afirmaciones que sólo contienen predicados monádicos y que *todos* los predicados monádicos tienen una estructura relacional oculta. No acepto estos puntos de vista pero, sin embargo,

encuentro que la doble reducción puede ser fructíferamente utilizada en una definición revisada de las relaciones externas:

Definición revisada: *Rab* es una relación externa si (i) es una relación externa de acuerdo con la definición tentativa anterior y *ya sea* (ii) el predicado reductor "F_i" G_j no tiene componentes relacionales ocultos *o* (iii) ninguna F_i con una estructura relacional oculta contiene una variable ligada, uno de cuyos valores posibles es *b*, y ninguna G_j, con una estructura relacional oculta, contiene una variable ligada, entre cuyos valores posibles está *a*.

Contrariamente, *Rab* es interna en uno de estos dos casos: o bien la relación no puede ser reducida o la reducción es tal que implícitamente nos referimos a *b* en alguna afirmación "$F_i a$" o a *a* en alguna afirmación "$G_j b$". Obsérvese la diferencia entre referir directamente a otro individuo, como en la afirmación "*Rab*" e indirectamente, al mismo individuo, como en la afirmación "$(\exists x)$ (Rax)" en donde la referencia a *b* puede darse a través del hecho de que *b* es un miembro de la clase de individuos de la cual puede obtenerse el valor de la variable *x*. Pienso que es intuitivamente razonable decir que la referencia a *b* es interna en ambos casos. Esto significa, por ejemplo, que no sólo la relación "poder sobre alguien" sino también la relación "más poder que" son relaciones internas, de acuerdo con la definición revisada. En un modelo simple, por ejemplo, la cantidad de poder se deriva del número de personas sobre quien se ejerce control en un paso o dos pasos.[30] (Aquí control significa lo mismo que "poder sobre alguien".) Esto quiere decir que el análisis de la afirmación "*a* tiene más poder que *b*" se da de la manera siguiente. Primero, escogemos algunos predicados reductores, por ejemplo, "*a* tiene dos grados de poder" y "*b* tiene un grado de poder". Observamos, entonces, que éstos pueden ser analizados como

"$(\exists x)$ $(\exists y)$ $(x \neq y \,\&\, Dax \,\&\, (Dxy \vee Day) \,\&\, (\forall z)$ $((Daz \vee Dxz \vee (Day \,\&\, Dyz)) \supset (z = x \vee z = y)))$"

y

$$(\exists x)\ (Dbx \,\&\, (\forall y)\ (Dby \supset y = x))",$$

respectivamente, donde "*Dxy*" significa "*x* tiene control sobre *y*". Si suponemos que *b* está entre los individuos sobre los que se ordenan las variables *x* e *y*, en la primera expresión, y *a* entre los individuos sobre los que se ordena la variable *x*, en la segunda, entonces, claramente "*a* tiene más poder que *b*", se ha demostrado que es una relación interna de acuerdo con la definición revisada. Por supuesto, es posible que se estén comparando individuos en sistemas sociales desarticulados, como en la afirmación "los reyes aztecas precolombinos tenían más poder que sus homólogos europeos", en cuyo caso la relación es externa, en ambas definiciones. Sostengo, sin embargo, que en la mayor parte de los casos en los que las comparaciones

de cantidad de poder parezcan pertinentes, estamos ocupándonos de comparaciones entre individuos pertenecientes a la misma red social.

Sustancialmente, el mismo análisis es válido para la relación entre (1) y (2) o entre (5) y (6). Este no es necesariamente el caso en que el predicado reductor "tener *n* grados de dinero" es en sí mismo reductible a la relación de explotación (como la relación "tener *n* grados de poder" era reductible a la relación "poder sobre alguien"), pero seguramente debe ser reductible a *alguna* afirmación relacional, que incluye variables ligadas, cuyo rango son los individuos y sus recursos. Pensar que la posesión de dinero es una propiedad natural del individuo a la par de su edad o altura, sería un caso de fetichismo similar al "fetichismo de la mercancía", atribuir al producto la propiedad de ser una mercancía, como si fuera una característica natural e inherente. La gente es rica porque otras personas aceptan su dinero; los productos son mercancías porque son tratados como mercancías en el intercambio con otras mercancías. La falacia empirista si es que hay alguna, no consistiría en la utilización de las relaciones externas en el análisis de clase sino, más bien, el descuido del carácter implícitamente racional de los atributos utilizados para las comparaciones. No se puede empezar describiendo individuos aislados para después ir hacia la definición de las relaciones (comparativas) entre ellos que constituyen la sociedad, porque las relaciones (interaccionales) deben estar implícitamente presentes desde el inicio.

La posesión de prestigio o estatus es quizá más obviamente un atributo socialmente definido. Para determinar la jerarquía y el prestigio objetivos de una sociedad, usualmente uno consigue primero los niveles subjetivos de una muestra lo suficientemente amplia y luego avanza para agregarlos, de alguna manera, al conjunto de niveles. En algunos casos "suficientemente amplio" significa, entre otras cosas, que todas las categorías, (por ejemplo, las ocupaciones) que están siendo ordenadas estarán ellas mismas representadas entre los que responden, de manera que será posible, también, estudiar la imagen del yo (*self*) en forma sistemática. Puede resultar, por ejemplo, que nadie ubica su propia ocupación al final de la lista y, sin embargo, habrá alguna ocupación que está al final de la lista total, de manera que en algunos casos (5) será verdadero y (6), falso. De nuevo sería incorrecto en nuestra definición revisada decir que (6) representa una relación externa si *b* es uno de los individuos cuyo ordenamiento subjetivo entra en la construcción de la escala agregada sobre la base de la cual adscribimos a *a* una cierta cantidad de prestigio.

Notas

1. Quine (1959) es una buena introducción a la lógica proposicional y cuantificacional en un nivel elemental. Church (1956) le da un tratamiento más avanzado. Son buenas introducciones a la lógica modal las de Snyder (1971) y Hughes y Cresswell (1972).

2. Como advierte Dummet (1973, págs. 8–9), los cuantificadores anidados están conectados de manera inextricable con el uso de las relaciones, en el sentido de que cada oración que contiene predicados de un lugar solamente es equivalente a alguna oración en la

que ningún cuantificador se sitúa dentro del campo de ningún otro cuantificador.

3. Davidson (1973, pág.717).

4. Marx (1845–46, pág. 33).

5. Marx (1867, págs. 484 y sigs.)

6. Las fórmulas con cuantificadores anidados pueden ser interpretadas de manera instructiva y divertida, como lo propuso por primera vez Jaakko Hintikka. Explotando el hecho (para esto véase Church, 1956, págs.209 y sigs.) de que cualquier oración cuantificadora puede ser convertida en *prenex normal form* (una cadena inicial de cuantificadores seguida por una matriz que no contiene ningún cuantificador), Hintikka (1973, pág.63) avanza a comparar la oración con un juego con la Naturaleza: "Mi fin en el juego es lograr una instancia sustitutiva de la matriz verdadera. El objetivo de mi Oponente (la Naturaleza) es hacer que el resultado del juego sea una falsa instancia sustitutiva de la matriz. Cada cuantificador existencial marca mi movimiento: elijo (produzco, encuentro) un individuo cuyo nombre queda sustituido por la correspondiente variable ligada. Cada cuantificador universal marca un movimiento de mi opositor: él es libre para producir un individuo cuyo nombre queda sustitituido por las variables ligadas al cuantificador universal. El orden de los movimientos corresponde al orden de los cuantificadores...Para que una oración sea verdadera es necesario y suficiente que yo tenga (de hecho) una estrategia ganadora en el juego correlativo, en el sentido de la teoría de juegos".

7. Los nombres "S–4" y "S–5" para estos sistemas se derivan de Lewis y Langford (1932). Los sistemas son estudiados extensamente en la literatura; véase para mayores detalles Snyder (1971).

8. Al reflexionar sobre su análisis de la lógica deóntica, Snyder (1972, pág.197) observa: "Esta excursión hacia la elaboración de un sistema de la lógica modal para encarar una necesidad particular debería indicar que mientras no sea el caso de que un sistema formal se elabora teniendo *cualquier* combinación de propiedades que podamos soñar, hay una gran flexibilidad en la maquinaria disponible y *pueden* desarrollarse sistemas formales que satisfagan condiciones más bien especializadas".

9. Véase especialmente Prior (1967).

10. Para una analogía entre las modalidades alética y deóntica véase Leibniz (1948, pág.606). Además de esta analogía formal, Leibniz postula también una conexión sustancial entre dos conjuntos de nociones, definiendo las acciones obligatorias como lo que es necesario para un buen hombre (Leibniz, 1948, pág. 606).

11. Véase para una discusión muy completa Snyder (1971, capítulo V).

12. Russell (1905).

13. Ejemplo (con modificaciones) tomado de Snyder (1971, pág. 204).

14. Hughes y Cresswell (1972, págs. 197–198) argumentan que es posible definir un sistema en el cual la fórmula '$N(\exists x)\ (Fx) \quad (\exists x)(N(Fx))$ resulte válida permitiendo que la variable se mueva entre los objetos *intensionales*, es decir, entre los conceptos individuales más bien que entre los individuos. Los objetos intensionales difieren de los objetos ordinarios por el criterio de identidad: como objetos ordinarios, Walter Scott y el autor de *Waverley* son idénticos pero, como objetos intensionales, son diferentes. En este sentido, hay un objeto intensional del cual es verdadero que es presidente de la república, por ejemplo, el objeto El Presidente de la República. Aun ignorando los problemas generales que surgen tan pronto admitimos los objetos intensionales, el interés de esta interpretación es severamente circunscripto por el hecho de que hace resultar válida no sólo la fórmula anterior de conjunto sino incluso la consecuente, en sí misma.

15. Esta frase abstrusa podría tener el siguiente significado. Supongamos que una población esta distribuida en una dimensión económica y, digamos, política de acuerdo con las distribuciones f y g. En el nivel estructural, podría interpretarse que la teoría marxista afirma que existe una función F tal que $g = F(f)$. En el nivel individual, el conocimiento de la posición de una persona en la dimensión económica debería dar la distribución de probabi-

lidad de su posición en otra dimensión.

16. Para esta tesis véase especialmente Ollman (1976). Estoy en deuda con Hayward Alker (h), por haber llamado mi atención sobre el tema.

17. Marx (1879–80, pág. 362).

18. Marx (1847, pág. 349).

19. Leibniz (1903, pág. 355).

20. Véase por ejemplo Russell (1910, pág. 374):"Yo sostengo que hay hechos tales que *x* tiene la relación *R* a *y*, y que tales hechos no son en general reductibles a o inferibles de un hecho *x* solamente ni de un hecho *y* solamente: ellos no implican que *x* e *y* tengan alguna complejidad o alguna propiedad intrínseca que los distinga de *z* y *w* las que no tienen la relación *R*. Esto es lo que yo significo cuando digo que las relaciones son externas". Esta caracterización de las relaciones externas sería, para mí, parte de la definición de las relaciones internas.

21. Véase Rorty (1967) para una buena exposición.

22. Esta aclaración se refiere a la posibilidad de que nuestras intuiciones lingüísticas puedan cambiar ellas mismas como resultado del análisis conceptual hasta que llegamos a lo que Rawls (1971, págs. 20 y sigs.) llamó un "equilibrio reflexivo".

23. Donne, *Woman's Constancy*. Estas líneas deben leerse a la luz de *The Broken Heart* y *The Paradox*, a los cuales se hace referencia en el capítulo 4. Para un punto de vista similar de otra autoridad sobre el amor, véase la siguiente elección de un confidente:"Quizá la cuestión más sabia es confiar en uno mismo. Utilizando nombres prestados pero incluyendo todos los detalles pertinentes escriba esta noche lo que ocurrió entre usted y su amante y los problemas con los cuales está enfrentado. Después de una semana, si está sufriendo de amor apasionado (*si vous avez l'amour–passion*) será enteramente otro y, entonces, al leer la historia de su caso, podrá darse a sí mismo buenos consejos" (Stendhal, *De l' Amour*, capítulo 34). ¿Quién está aconsejando a quién, aquí?

24. Moore (1922).

25. Mates (1968); Ishiguro (1972a,1972b); Hintikka (1972).

26. En realidad, intenta operar tales reducciones en las afirmaciones relacionales "Pedro es como Pablo","Tito es más sabio que Cayo" y "Paris ama a Elena"; ver para mayores detalles Mates (1968)y Hintikka (1972). Opuesto a las definiciones aquí propuestas, Leibniz admite en sus reducciones conectivas funcionales no veritativas.

27. Dice, por ejemplo, que"no hay ningún término tan absoluto ni tan desprendido que no contenga relaciones, y cuyo análisis perfecto no conduzca a otras cosas, incluso a todas las otras cosas" (Leibniz, 1875–90, vol. V, pág. 211).

28. Un modelo simple propuesto por Kemeny y otros (1966, pág. 384) en el cual "el poder sobre alguien" es la noción básica, cantidades absolutas de poder, una noción derivada y "más poder que" derivada en una propuesta ulterior. Un modelo más complejo es el dado por Coleman (1973a), en el cual las nociones básicas son el control de los actores sobre los recursos y los intereses en los recursos, de los cuales se puede derivar, primero, la relación "poder sobre alguien" y cantidades absolutas de poder y, segundo, la relación "más poder que".

29. Hintikka (1972).

30. Este es el modelo de Kemeny y cols. (1966).

2

Razonamiento posibilístico en las ciencias sociales

En este capítulo examinaré algunos modelos de las ciencias sociales que en diversas formas exhiben "características posibilísticas", por ejemplo, modelos en los que las posibilidades puras (como diferentes de las probabilidades cuantificadas) tienen una función teórica importante. En algunos casos intentaré reflexionar sobre las articulaciones eventuales entre estos modelos y los conceptos modales, detalladamente explicados en el capítulo 1. En otros casos, la articulación es, cuando mucho, indirecta. La función de este capítulo no es tanto ofrecer aplicaciones específicas de la lógica modal (esta es una tarea de posteriores capítulos), sino convencer al lector escéptico de que la noción general de mundos posibles tiene ya un lugar en las ciencias sociales. Los modelos presentados son los siguientes:

–movilidad social,
–gramáticas de estados finitos,
–sistemas de parentesco,
–evaluación del ingreso nacional neto,
–cambios endógenos de los gustos.

Movilidad social

El modelo simple para el estudio de movilidad social es el enfoque en cadena de Markov. Presumiendo que tenemos n clases, el enfoque postula que para cada par de clases i,j, hay un número p_{ij} que caracteriza únicamente la probabilidad de un individuo en la clase i en un tiempo t, estando en la clase j en el tiempo $t+1$, siendo este número independiente (i) de la historia anterior de la movilidad del individuo, (ii) de otras características reales diferentes de la pertenencia de clase y (iii) de los eventos en el tiempo calendario. En principio, la misma conceptualización puede ser utilizada para el estudio de la movilidad entre generaciones.[1] En muchos casos, una

matriz de transición de estos p_{ij} contendrá algunos ceros, o sea, para algunos pares de clases hay cero probabilidad de movilidad (un paso). Esto no significa que es imposible para los descendientes de un miembro de la primera clase pertenecer, finalmente, a la segunda. Puede llegar a ser posible hacer la transición en diversos pasos, cada uno de los cuales está representado por una entrada no nula en la matriz de transición. Para estudiar este problema podemos reemplazar la matriz probable p_{ij} con la matriz posibilística d_{ij}, donde $d_{ij} = p_{ij}$ si $p_{ij} = 0$ y $d_{ij} = 1$, otra vez. Llamamos *matriz de adyacencia* a la matriz d_{ij}.[2] Expresa de manera simplificada la esencia de las posibilidades de movilidad de corta duración. Es muy posible que para algunos objetivos esta matriz simplificada sea más fecunda analíticamente que la estructura probabilística completa. Puede darse el caso, por ejemplo de que posibilidades *percibidas* de movilidad tengan la forma simple Sí–No antes bien que el carácter gradual de la matriz p_{ij}. El único negro que llega a ser juez de la Corte Suprema o la mujer que llega a ser presidente de una gran empresa pueden tener un impacto cualitativo bastante independiente de las realidades numéricas subyacentes.[3]

Habiendo construido la matriz de adyacencia, podemos ir a la definición de matriz de *alcanzabilidad*, cuyas entradas dan la posibilidad de ir de la clase *i* a la *j*, en cualquier número de pasos.[4] Si la matriz de alcanzabilidad tiene algunas entradas nulas, esto significa que la estructura social en ciertos aspectos se asemeja a un sistema de castas, siendo algunas posiciones inaccesibles para los miembros que comienzan en otras posiciones. Tanto la relación de adyacencia de corto plazo como la relación de alcanzabilidad de largo plazo pueden ser interpretadas en términos de la lógica modal. Para un estado dado del mundo, por ejemplo, una determinada distribución de individuos en las clases, los estados accesibles son los estados posibles, de acuerdo ya sea con la matriz de corto plazo o con la de largo plazo. Los conceptos modales de posibilidad y necesidad se siguen entonces de inmediato. En particular, podemos anotar que ahora se hace posible hablar de *rasgos necesarios* de un individuo, es decir, los rasgos que caracterizan a todos los individuos de todas las clases a las que pueden acceder en un paso (o en un número de pasos). Algunos individuos pueden ser contingentemente *F*, o sea, pueden tener acceso inmediato o final a las clases no *F*, mientras otros pueden ser necesariamente *F*. Esta distinción puede concebirse como correlacionada con otras propiedades de los individuos. No digo que siempre lo estén, sólo que esta interpretación modal de la movilidad produce inmediatamente esta noción de rasgo necesario. Si uno cree, como yo, que es bueno tener tantas y tan variadas estructuras generadoras de conceptos como sea posible, este es, por lo menos, un argumento *prima facie* para que el estudio de estas ideas lógicas valga la pena.

Puede observarse que los *rasgos necesarios* no son lo mismo que *restricciones* : el primero es un concepto más amplio. Para explorar esta distinción podemos hacer una corta digresión desde el tema de esta

sección y considerar la idea de un conjunto factible en la teoría económica, tal como un conjunto que es definido por un número de restricciones. Por ejemplo, dado un estado económico inicial, pueden imponerse diversas condiciones para limitar el conjunto de estados que pueden seguir al estado inicial. Además de las restricciones tecnológicas, puede no permitirse que el consumo caiga más allá de cierto límite mientras la inversión debe mantenerse por encima de un nivel mínimo. Puede resultar que todas las vías para satisfacer estas restricciones (y quizás otras, adicionales) simultáneamente comprenden la utilización de tecnología altamente contaminante, la cual se torna, entonces, un rasgo necesario, sin ser una restricción explícita. Un tipo de ejemplo muy distinto puede ser el siguiente. Supongamos que en un conflicto relacionado con la reforma educacional tanto los maestros como los padres sostienen algunos reclamos que actúan como restricciones a la Dirección de Educación. Los maestros piden menos horas de enseñanza y los padres más matemáticas y ciencia, de lo cual se deduce que todos los resultados factibles incluyen una reducción de lenguas e historia. Este rasgo, siendo necesario, no es impuesto por ninguno como una restricción y puede, en verdad, ir o estar en contra (levemente) de las preferencias de todos. La distinción entre restricciones y rasgos necesarios puede ser también comparada con la distinción entre axiomas y teoremas en un sistema formal. Cada axioma puede ser visto como un teorema, con el rasgo particular de que es derivado solamente de un axioma, por ejemplo, de sí mismo. Los otros teoremas son derivados de varios axiomas tomados de conjunto. Tomados por sí solos, los axiomas son restricciones. Tomados conjuntamente, generan características necesarias.

Volviendo al estudio de la movilidad social, ha resultado que el procedimiento de Markov no es realmente adecuado para este fin.[5] El principal hallazgo empírico que lo invalidó fue el hecho de que la proporción de individuos que permanecen en una clase desde un tiempo t a un tiempo $t + 2$ es mucho mayor de lo que podría esperarse sobre la base de la presunción de Markov.[6] Mientras más tiempo un individuo ha sido observado permaneciendo en una clase dada, mayores son las posibilidades de que se mantenga allí en el siguiente período. Como siempre, en dichos casos, esta regularidad empírica puede ser interpretada teóricamente de dos maneras diferentes. O bien estamos tratando con un efecto de muestreo, de manera tal que el individuo se revela a sí mismo como siendo (y habiendo sido siempre) un típico "permanecedor" más que un "movedor", o estamos ante un efecto posterior, en el que el individuo mismo es cambiado por una prolongada pertenencia de clase.[7] La última noción —que ha sido llamada hipótesis de inercia acumulativa— es especialmente interesante en este contexto. Ella implica que el "halo de posibilidades" que circunda lo efectivo puede cambiar por el mismo hecho de que el tiempo avanza sin *ningún* cambio en el estado real. Esta es una idea que será nuevamente explorada en el capítulo 3. Aquí sólo señalaré que no estoy implicando la idea absurda de que las posibilidades de movilidad pueden

cambiar sin *ningún* cambio en el mundo real. La cuestión es simplemente que las probabilidades (y posibilidades) de transición pueden cambiar sin ningún cambio en el individuo *como se le definió*, por ejemplo, sin ningún cambio en la variable de actual pertenencia de clase. Otras cosas, seguramente, han cambiado, por ejemplo, el nivel de aspiración y el ingreso, relativo o absoluto, etcétera.

Más formalmente: dada la observación empírica de que el valor de la variable x en un tiempo $t + 1$ no puede predecirse a partir del valor de x a t, podemos extender el modelo en una de dos vías: o bien podemos introducir el valor t de algunas variables nuevas *y,z*...o podemos introducir el valor t-1, t-2...de x. Desde el punto de vista metafísico, el último proceso no es muy satisfactorio puesto que parece implicar que el pasado puede a veces causar una influencia causal sobre el presente, por encima de la influencia que está mediada por las huellas dejadas por el pasado en el presente.[8] Por otra parte, puede ser más conveniente utilizar valores pasados de la misma variable, más bien que introducir nuevas variables. Aquí puede haber un intercambio entre la elegancia explicativa y el realismo metafísico.[9]

Gramáticas de estados finitos

También se puede usar la relación entre el modelo probabilístico de Markov y la correspondiente estructura posibilística para iluminar nuestra comprensión del proceso lingüístico. Queremos contestar la siguiente pregunta: ¿Cuáles son las reglas que gobiernan la formación de oraciones en un lenguaje dado? Más precisamente, esta pregunta puede ser entendida de dos diferentes maneras: como una pregunta empírica acerca de las oraciones pronunciadas por los hablantes nativos en la vida cotidiana y como una investigación teórica sobre las oraciones que los hablantes nativos juzgan gramaticales. Para empezar con la primera cuestión, es suficientemente claro (i) que las palabras son pronunciadas en diferentes frecuencias y (ii) que una palabra dada es seguida por otras palabras de frecuencias variables, siendo cero la frecuencia para muchas o la mayoría de las siguientes. Un primer intento para entender la formación de la oración sería la siguiente hipótesis: la probabilidad p de que una oración "$a_1 ... a_n s$" sea pronunciada, en la cual las a_i son palabras sueltas y "s" representa "fin de la oración", es el producto de las probabilidades

$$p_1 \cdot p_{1,2} \cdot p_{2,3} \cdots p_{n-1,n} \cdot p_{n,s}$$

donde p_1 es la frecuencia de a_1 y $p_{i,j}$ es la probabilidad de a_j, dándose inmediatamente después de a_i. Una reflexión rápida, sin embargo, demuestra que éste no es un modelo muy plausible. La probabilidad de que a_j ocurra después de a_i depende fuertemente de las palabras que preceden a a_i en la secuencia. Una sucesión de palabras pares compatibles pueden dar un todo incomprensible, como en el siguiente ejemplo: "La cabeza y en un

ataque frontal a un escritor inglés que el carácter de este punto es por lo tanto otro método para las letras que el tiempo de quien habló del problema para un inesperado."[10]

Tomando en cuenta la historia anterior, si se permite la probabilidad de que una palabra aparezca dependiendo de, digamos, las cuatro palabras que la anteceden en la secuencia, no se garantizan enunciados probables o,incluso, posibles. Utilizando grupos de cinco palabras con una frecuencia de palabras representativa del inglés, un conocido estudio [11] generó lo siguiente: *Road in the country was insane especially in dreary rooms where they have some books to buy for studying Greek.* En principio, uno podría intentar superar esta dificultad haciendo que la probabilidad de aparición de una palabra dependa en un número mayor de palabras antecedentes, pero el número requerido de probables transiciones rápidamente se haría imposiblemente grande. Además, hay un problema que concierne a la relación entre la probabilidad de un enunciado y la gramaticalidad de una oración. Si se comienza con un orden superior del modelo de Markov (por ejemplo, un modelo de Markov capaz de tomar en cuenta la historia anterior) para el estudio probabilístico de enunciación de la oración, entonces sería bueno que la versión posibilística de este modelo generara exactamente el conjunto de todas las oraciones gramaticales. "Gramatical" simplemente significaría entonces, "tener no cero probabilidad de ser pronunciada por un hablante nativo".[12] Puede demostrarse, sin embargo (y lo haré en un momento), que ningún tipo de modelo markoviano puede generar todas las oraciones que un hablante nativo aceptaría como gramaticales y sólo aquellas oraciones. Puede incluso hacerse la afirmación más fuerte de que algunas oraciones gramaticales tienen probabilidad cero de ser pronunciadas si, por ejemplo, el tiempo requerido para su expresión excede la duración del mundo físico. Ciertamente hay una necesidad de un modelo probabilístico de enunciación de una oración, pero nunca puede servir como fundamento para un modelo posibilístico de gramaticalidad de la oración. En este punto hay un contraste agudo con el modelo discutido en la sección previa, en el cual la posibilidad fue definida directamente como probabilidad no cero.

El argumento contra un modelo de gramaticalidad markoviano [13] utiliza la noción de un *estado finito de la máquina*; los diversos modelos antes descritos son casos especiales de éste. Existiendo el estado inicial, la máquina tiene n estados $s_1 \ldots s_n$ s_1. Cuando la máquina se encuentra en estado s_i hay un conjunto bien definido S_i de estados a los cuales puede pasar desde s_i. Cuando la máquina vaya de s_i a $s_j \varepsilon\ S_i$, puede escribir cualquier palabra de un conjunto de palabras S_{ij}. Una oración gramatical es una cadena de palabras que pueden ser el resultado de la ida de la máquina de un estado inicial vía otros estados y de regreso al estado inicial. Para la generación de cadenas más largas que n (que contengan más que n palabras), la máquina tiene que pasar, evidentemente, a través de una secuencia de estados en los que por lo menos un estado, digamos s_k, ocurra dos veces. Entre las dos ocurrencias de s_k, tiene que haber una secuencia

de estados $s_{n1}\, s_{n2} \ldots s_{np}$, con una subcadena escrita. Pero si $s_k\, s_{n1} \ldots s_{np}\, s_k$ es una secuencia posible de estados, entonces también lo es

$$s_k\, s_{n1} \ldots s_{np}\, s_k\, s_{n1} \ldots s_{np}\, s_k,$$

con la repetición de la subcadena de palabras. Esto significa también que la oración en cuestión debe contener una subcadena que pueda ser repetida (por extensión del mismo argumento) indefinidamente, cualquier cantidad de veces, sin violación de la gramaticalidad. Para resumir, si se va a generar el lenguaje por una gramática de estado finito, cualquier oración suficientemente larga debe contener una subcadena que pueda ser repetida cualquier cantidad de veces, sin violación de la gramaticalidad.

Demostraremos ahora que éste no es el caso en las oraciones de la forma

A_k: Si^k llueve (entonces ventea)^k, $k=1,2 \ldots$,

en donde los exponentes significan repetición. Es claro que todas las A_k son perfectamente gramaticales. También está claro que para cualquier gramática de estados finitos con n estados propuestos, podemos encontrar una k tal que la longitud de A_k exceda n. Esto significa que para una k lo suficientemente grande, A_k debe contener una subcadena que pueda repetirse indefinidamente. Sin embargo, es una cuestión sencilla chequear que esto no puede ser el caso para ninguna de las subcadenas de una A_k. Quiere decir que hemos encontrado una oración que (i) en el campo intuitivo tiene que ser aceptada como gramatical, (ii) no puede ser generada por la gramática de estados finitos en cuestión. Como dichos contraejemplos pueden encontrarse para cualquier gramática de estados finitos (aun si no hay contraejemplo útil contra todas estas gramáticas), éstos son inadecuados, en principio, para el análisis de los lenguajes naturales. (En un apéndice del presente capítulo se prueba un teorema similar para una categoría más amplia de gramáticas.) En la raíz de este análisis posibilístico hay un postulado básico de la lingüística moderna, la idea de que lo que estudiamos es competencia más bien que ejecución, lo posible más bien que lo real.

Los sistemas de parentesco y el pensamiento estructuralista

Hay una escuela de pensamiento cuyo objeto de estudio *par excellence* es lo posible en vez de lo real: la escuela francesa de estructuralismo. No hay acuerdo en general acerca de la esencia ni siquiera de la existencia, de este grupo; algunos de sus presuntos miembros niegan dicha afiliación. Sin embargo, pienso que estaremos en territorio firme si escogemos a Claude Lévi-Strauss como el principal exponente del pensamiento estructuralista.

Su procedimiento general puede ser brevemente descrito como finitista, combinatorio y posibilístico. Rasgos similares caracterizan los trabajos de pensadores como Roland Barthes, Jaques Lacan y Louis Althusser (antes de su autocrítica). Me referiré brevemente a la escuela marxista de Althusser pero, fuera de esto, me limitaré a Lévi-Strauss.

Es posible seguir diversas etapas en las reflexiones metodológicas de Lévi-Strauss. Una primera diferenciación entre modelos mecánicos y estadísticos fue propuesta en 1958.[14] Estos conceptos quedan cuidadosamente sin definir y no hay consenso en cuanto a su significado. [15] Un aspecto importante, sin embargo, parece suficientemente claro: un modelo mecánico describe a la sociedad por medio de un conjunto de reglas, de manera que la hace aparecer como uno de muchos conjuntos posibles, mientras que un modelo estadístico es sólo un resumen empírico de la sociedad real. Este aspecto se clarifica en una diferenciación introducida en un ensayo de 1964 sobre la diferencia entre las ciencias humanas y las ciencias sociales.[16] Las primeras incluyen disciplinas como la lingüística (*la* ciencia humana *par excellence*), crítica de textos, psicología y antropología social (*a la* Lévi-Strauss), mientras que las últimas incluyen economía, sociología y ciencias políticas. Las ciencias sociales están muy articuladas a una sociedad determinada, para la que llevan a cabo funciones técnicas y quizás ideológicas. Son esencialmente limitadas y prácticas. De otro lado, las ciencias humanas son generales y teóricas, estando, en este aspecto, mucho más cerca de las ciencias naturales y exactas, que claramente son el ideal metodológico de Lévi-Strauss.[17] El objeto de las ciencias humanas es la sociedad como tal o el hombre como tal; [18] el alcance de la variación de todos los objetos posibles de una clase determinada; [19] los rasgos comunes de todos los objetos posibles de una clase determinada; [20] lo posible y lo necesario, nunca sólo lo real.

La analogía con la lingüística y especialmente con la fonología es omnipresente en el trabajo de Lévi-Strauss, utilizada por él para justificar su método combinatorio en la antropología. Es una práctica común en las ciencias sociales (y las humanas) trabajar con variables dicotomizadas, cuyas combinaciones nos permiten construir una clasificación aproximada, pero alentadoramente fértil. En fonología (y especialmente en el trabajo de Roman Jakobson), sin embargo, las oposiciones binarias son vistas como rasgos de la realidad más bien que como dispositivos heurísticos. Se dice que el análisis de los fonemas en términos de ausencia o presencia de unos 12 rasgos distintivos no es sólo una descripción posible entre muchas otras, sino que corresponde a la estructura mental del hablante.[21] Si se acepta esto, la utilización del análisis combinatorio para generar todo posible sonido, por así decirlo, es claramente justificada. Si se rechaza, el resultado de la combinatoria no tendrá ningún interés o significado particular, porque cualquiera de las otras conceptualizaciones igualmente válidas podría ser utilizada para generar un conjunto diferente de posibles. No es necesario discutir acá si la afirmación se justifica en la lingüística. Lo que me parece bastante cierto es que las afirmaciones hechas por Lévi-Strauss cuando analiza la cocina en "gustemas"[22] o los mitos en "mitemas"[23] no pueden darse

por buenas. Es verdaderamente divertido leer una descripción de la diferencia entre la cocina francesa y la inglesa en términos de tres oposiciones binarias, pero no se ofrece ninguna prueba de que estas oposiciones existan en ningún lugar fuera de la mente de Lévi-Strauss. En mi opinión su trabajo posterior sobre esquemas de clasificación y mitología degeneró en un ejercicio arbitrario y gratuito de ingenio desubicado, totalmente falto de validez intersubjetiva.

Sin embargo, muchos toman en serio este trabajo, hecho éste que pide explicación. Una razón puede ser el prestigio derivado de sus primeros trabajos acerca de los sistemas de parentesco, en los cuales el análisis combinatorio parece, en verdad, haber ofrecido una comprensión fundamental. Incluso en este caso pienso que algunos de sus análisis están seriamente debilitados por un razonamiento arbitrario, como su utilización del "átomo de parentesco", que consiste en las relaciones entre hermano y hermana, marido y mujer, padre e hijo, tío materno y sobrino. La tesis general propuesta[24] fue que, de estas cuatro relaciones, dos serían siempre negativas y dos positivas. A la crítica evidente de que las relaciones humanas son típicamente ambivalentes, siendo al mismo tiempo (pero en aspectos diferentes) positivas y negativas, Lévi-Strauss ofrece una respuesta [25] que, hasta donde yo entiendo, quiere decir que cualquier *diferencia* puede servir como una *oposición*. Esto, sin embargo, parecería dejar al antropólogo total libertad de elección en cuanto a qué relaciones serán llamadas positivas y cuáles, negativas, haciendo la teoría tan flexible como para que sea imposible refutarla.

El análisis más influyente de Lévi-Strauss sobre parentesco está, en realidad, en *Las estructuras elementales del parentesco* donde (junto con el matemático André Weil) demostraba que dados ciertos requerimientos concernientes al matrimonio y la descendencia hay sólo un número limitado (y pequeño) de sistemas de parentesco posibles. Posteriormente, otros escritores han ofrecido análisis más elaborados.[26] En esta exposición definiré un *sistema de parentesco* como cualquier división de la sociedad en "tipos" tales que (i) sólo individuos del mismo tipo pueden casarse entre sí, (ii) el tipo de un individuo depende sólo del sexo del individuo y del tipo de sus padres (iii) dos chicos (o chicas) que tengan padres de distin-to tipo serán, ellos mismos, de diferentes tipos. El tercer requerimiento implica que el cambio en el tipo de los padres a los hijos será una permutación del conjunto de tipos y, de manera similar, el cambio de los tipos de los padres a los hijos. Con tipos n hay $n!$ permutaciones por cada sexo y, por lo tanto, $(n!)^2$ sistemas diferentes de parentesco. Este número aumenta tan rápidamente con n que con sólo un número moderado de tipos hay demasiados sistemas de parentesco como para que podamos revisarlos a todos. Por lo tanto, cuatro tipos implican 576 sistemas de parentesco; de éstos, un número importante serán isomórficos con respecto a otros pero, aun así, muchos permanecerán. Imponiendo algunas condiciones más, sin embargo, es posible reducir drásticamente el número de sistemas de paren-

tesco. Llamenos *sistemas de parentesco permisibles* a los sistemas que satisfagan tanto a las mencionadas (i), (ii) y (iii), como (iv) a la norma de que depende sólo de la clase de relación que a un hombre se le permita casarse con una pariente femenina de un tipo dado; (v) en particular, a ningún hombre se le permite casarse con su hermana y (vi) es posible para dos individuos cualesquiera que algunos de sus descendientes se casen entre ellos. Utilizando el razonamiento teórico de matriz y de grupo no es difícil comprobar que hay más bien pocos sistemas de parentesco permisibles; por lo tanto, con cuatro tipos sólo hay seis maneras no isomórficas de satisfacer (i)-(vi).

¿Cuáles son las consecuencias de estos resultados? En el presente contexto yo enfatizaría el hecho de que nos permiten ubicar una estructura de parentesco dada en el campo de todas las estructuras posibles de parentesco, de la misma dimensión (con el mismo número de tipos). Aún más importante es el hecho de que este campo es no sólo de cardinalidad limitada sino típicamente muy pequeño, de manera tal que permite ver si se da de hecho cualquier otra de las estructuras posibles en el campo. Entre los seis sistemas permisibles en el campo de dimensión 4, sólo dos resultan darse en las sociedades reales: una pregunta obvia es, entonces, cuáles son los rasgos de los otros sistemas que explican su ausencia. La respuesta a esta pregunta también proveería, por consiguiente, una valiosa visión de la naturaleza de los sistemas que se dan de hecho en las sociedades reales. Aquí tenemos entonces otro sentido (metafórico) en el cual el halo de posibilidades que circunda lo real debería explorarse para lograr un entendimiento pleno de lo real, en sí mismo.[27]

No mantengo ninguna reyerta con este análisis, al contrario (para mi limitada competencia) aparece como una pieza ejemplar de razonamiento. Quisiera insistir, sin embargo, en los peligros propios –y los obstáculos– de la transferencia de estos métodos a otros campos. La naturaleza binaria de las relaciones de parentesco (el matrimonio y la descendencia) es parte del orden de las cosas y no un artefacto analítico. Lo mismo no puede decirse de, digamos, la relación entre las fuerzas de producción y las relaciones de producción en la teoría marxista. Analizar esta relación en términos de oposiciones binarias y luego pasar a generar todos los modos de producción posibles es un procedimiento totalmente arbitrario, basado en un razonamiento por analogía, falaz y pernicioso. [28] Como los gustemas de Lévi-Strauss, la tabla periódica de los modos de producción es seudociencia, del mismo orden que, digamos, la sociología inspirada en la biología de los años 1890.

Evaluación del ingreso nacional neto

El título de esta sección es tomado de un importante trabajo de Paul Samuelson donde, entre otras cosas, exploró la siguiente cuestión: ¿por qué querrían los economistas incluir productos tan antieconómicos como los

bélicos en el cálculo del ingreso nacional? Su respuesta es que "presumiblemente sirven como algún tipo de índice de cosas útiles que pueden ser producidas en tiempos mejores",[29] ejemplo de una respuesta claramente posibilística o contrafáctica. Otra cuestión –el problema central del trabajo– hace a las dificultades para comparar los niveles de bienestar social de dos estados distintos de la economía. Discutiré primero este problema dentro del marco de referencia de Samuelson, en el que se supone que dos estados incluyen el mismo conjunto (inalterable) de individuos. Luego añadiré observaciones acerca de la forma en la que A. K. Sen extendió recientemente este análisis para tratar las comparaciones entre los niveles de bienestar social de distintos grupos de personas.

Dados, entonces, dos estados que incluyen a los mismos individuos, el método más simple para comparar los niveles de bienestar social es el criterio de *superioridad de Pareto*: el estado *A* es mejor que el estado *B*, si los niveles de bienestar social de todas las personas son, por lo menos, igualmente grandes en *A* y en *B*, y el nivel de, por lo menos una persona, es estrictamente mayor en *A* que en *B*. Este criterio está abierto al comentario obvio de que, incluso aceptando que es correcto cuando es aplicable (lo cual es discutible, como se demuestra más abajo), es aplicable sólo a un subconjunto muy pequeño del total de números de pares de los estados económicos. El ordenamiento que resulta del criterio de superioridad de Pareto es, en verdad, parcial y no es de sorprenderse que los economistas hayan buscado extenderlo. El relato de estas extensiones no exitosas se ha realizado frecuentemente y hablaré brevemente al respecto. El *criterio de Kaldor o test de compensación* implica que el estado *A* es mejor que el estado *B* si, por medio de la redistribución del ingreso (un esquema de compensación) en *A* es posible hacer *A* a la Pareto superior a *B*. T. Scitovsky demostró, sin embargo, que el criterio de Kaldor no tiene la propiedad deseada de la antisimetría: de acuerdo con este criterio *A* podría ser mejor que *B* y *B*, también mejor que *A*.[30] Propuso, entonces, el *criterio de Scitovsky*, de acuerdo con el cual *A* es mejor que *B* si *A* puede, redistribuyendo, hacerse superior a la Pareto a B, pero no viceversa; sin embargo, este criterio resulta carecer de la deseada propiedad de transitividad.[31]

El artículo de Samuelson debe ser visto en el contexto de estos fracasos de extensión del criterio de Pareto. Intenta hacer explícitas las ideas posibilísticas incluidas tanto en los criterios de Kaldor como de Scitovsky, y el resultado es una contrafactibilización al por mayor de las economías reales. Bosquejaré brevemente sus ideas fundamentales para un mundo de dos personas y dos bienes. Suponemos que la distribución real de los bienes *A* y *B* entre los individuos I y II se da en las cantidades A_I, B_I, A_{II}, B_{II}, llevando a los niveles ordinales de utilidad (u_I y u_{II}), respectivamente. Un primer concepto introducido por Samuelson es la *curva de posibilidad de utilidad*: la frontera óptima de Pareto del conjunto de puntos u'_I, u'_{II} que puede lograrse por medio de la redistribución del total de las cantidades $A_I + A_{II}$, $B_I + B_{II}$ de los bienes *A* y *B* entre I y II. Argumentando según esta suposición de cantidades totales constantes de ambos bienes,

Samuelson postula que el único caso en el que, sin ambigüedad, puede decirse que una situación es mejor que otra es cuando la curva de posibilidad de utilidad de la primera cae totalmente por fuera de la curva de posibilidad de utilidad de la segunda (si marcamos u_I y u_{II} a lo largo de estos dos ejes). O sea que, incluso cuando A es superior a la Pareto a B, no podemos decir, sin ambigüedades, que A es mejor que B, puesto que si las curvas de posibilidad de utilidad que van a través de A y B se intersectan, redistribuciones simultáneas en ambos estados pueden hacer que B resulte superior a la Pareto a A.

El siguiente paso de Samuelson es introducir en escena la producción y la tecnología, de manera que no podemos seguir suponiendo que A y B estén presentes en cantidades totales constantes. El conocimiento de la función de la producción nos dirá qué posibles combinaciones (A', B') podrían haberse producido en cambio de ($A_I + A_{II}$, $B_I + B_{II}$) dado el factor de dotación existente. Para cada combinación tal, entonces, se da el caso de que podría haber sido distribuida entre I y II de un gran número de maneras. Esta doble contrafactibilización nos permite definir una noción más rigurosa de la curva de posibilidad de utilidad, por medio de los siguientes procedimientos. Primero, utilizamos nuestro conocimiento de la producción de posibilidades para determinar todas las posibles combinaciones de A y B que podrían haberse producido. Para cada combinación construimos una curva de posibilidad de utilidad en el sentido definido anteriormente. Finalmente, definimos la *curva inclusiva de posibilidad de utilidad* como el envoltorio exterior de todas las funciones particulares elaboradas de esta manera. Un estado es, entonces, mejor que otro, sin ambigüedad (y es mayor su ingreso nacional real), si y sólo si la curva inclusiva de posibilidad de utilidad del primero yace totalmente por fuera de la del segundo, por ejemplo, si todos los estados obtenibles a partir del segundo, a través de la reorientación de las actividades productivas y de la redistribución del ingreso, son también obtenibles del primero, pero no viceversa. De acuerdo con esta definición se nos permite decir que una economía que produce gran cantidad de bienes bélicos mientras causa hambre en el consumidor tiene un ingreso nacional mayor que una economía donde todos los consumidores son prósperos.

Me parece que el punto fundamental de Samuelson es el siguiente. *O bien* comparamos el ingreso nacional y los niveles de bienestar social en términos de la carencia *real* de satisfacción y sólo entonces el criterio de Pareto puede ser útil. *O* podemos pensar que la *potencialidad* de la carencia de satisfacción es, también, pertinente pero, entonces, tenemos que ir hasta las últimas consecuencias y contrafactibilizar todo, la producción y la distribución, en ambos estados. La contrafactibilización de una sola situación, la producción o la distribución, es como una vivienda a medio construir y no tiene justificación teórica. Estas no son afirmaciones obviamente verdaderas. Por el lado de la producción, la reorientación de las actividades productivas no es instantánea ni carece de fricciones. Si afirmamos que la sociedad "podría" hoy haber producido (A',B') en vez de (A,B)

tendríamos que agregar una aclaración tal como "si ciertas elecciones fundamentales en el pasado hubieran sido diferentes". No es evidente que la contrafactibilización del pasado (quizá bastante distante) sea pertinente cuando se la compara con los ingresos nacionales *de hoy*. Por el lado del consumo, Samuelson mismo enfatiza que no todos los esquemas de redistribución pueden ser políticamente factibles. El problema es que la única técnica de redistribución que no tiene un impacto sobre los incentivos laborales (y por lo tanto, sobre el total que debe ser redistribuido) es la así llamada "solución al bulto", que es virtualmente imposible hacer funcionar en la práctica.[32] Las técnicas que *son* posibles, tales como el impuesto progresivo, o el racionamiento conducirán, por otra parte, a la reducción de la eficiencia con un monto menor total para la redistribución. Por este motivo, Samuelson sugiere que estrechemos la curva inclusiva de posibilidad de utilidad para que comprenda sólo las posiciones obtenibles a través de medios *políticamente posibles* de redistribución del ingreso. Esta noción de posibilidad política es el tema del capítulo 3.

Según Kenneth Arrow y Tibor Scitovsky, "el método purista de Samuelson no fue totalmente aceptado por la profesión",[33] posiblemente por el mareo generado por los contrafácticos requeridos. En esta cuestión, sin embargo, el análisis de Samuelson es la sobriedad misma cuando se lo compara con el análisis reciente hecho por Amartya K. Sen, de la lógica de las comparaciones que involucran diferentes conjuntos de individuos.[34] Sen comienza haciendo la observación de que, aun en el caso de una comparación entre los estados A y B que involucre al *mismo* conjunto de individuos, hay una diferencia importante entre

(1) el grupo está mejor con A que lo que hubiera estado con B; y
(2) el grupo está mejor con A que lo que estuvo con B,

porque la estructura ordinal o cardinal de la relación de preferencia social puede cambiar con el tiempo. Una diferenciación similar es útil para comparar los niveles de bienestar social de dos grupos o comunidades diferentes de individuos. Suponiendo, para simplificar, que la comunidad tiene el mismo número de miembros [35] deberíamos diferenciar entre

(3) la comunidad I está mejor con A que lo que hubiera estado con B; y
(4) la comunidad I está mejor con A que la comunidad II con B.

Queremos discutir la afirmación (4), pero un análisis directo es claramente imposible si aceptamos la no aditividad de las utilidades individuales. Como sustituto, podemos ceñirnos a la afirmación contrafáctica (3) pero, entonces, debemos resolver dos problemas. Primero, tenemos que decidir qué personas de la comunidad I recibirán qué canastas de consumo de parte de la comunidad II. Sen menciona varias formas de enfrentar esta dificultad. Las dos más atractivas son las siguientes. Primera, que intentemos escoger alguna correspondencia privilegiada uno a uno entre perso-

nas de I y canastas de II, y digamos que (3) es verdadero si y sólo si I es socialmente preferido a II bajo esta correspondencia. Podríamos por ejemplo requerir que la canasta del individuo en el lugar *i* en la jerarquía de ingresos de II fuera dada al individuo de la misma jerarquía de I, o podríamos imponer la exigencia, mucho más fuerte, de que la afirmación (3) es verdadera si y sólo si I es socialmente preferible a II para *todas* las posibles *n*! correlaciones entre canastas e individuos. Segunda, tenemos que decidir qué significa "socialmente preferible". Esto, por supuesto, nos retrotrae a las cuestiones discutidas en los primeros párrafos de esta sección. No discutiré la solución de Sen, salvo señalar que no enfrentó el problema de la producción y la distribución contrafáctica, que es crucial en el trabajo de Samuelson. Parecería que en el método "purista", tendría que estar comprendida una comparación entre los niveles de bienestar social de dos comunidades distintas, (i) una contrafactibilización de la producción, (ii) una contrafactibilización de la distribución y, (iii) para cada par de contrafactibilizaciones sucesivas de la producción y la distribución, una imputación contrafáctica (o un conjunto de imputaciones) de canastas a personas.

En el capítulo 6 veremos surgir estas cuestiones de una manera muy natural en el análisis de problemas históricos tales como el impacto de la esclavitud en el Sur *ante-bellum* sobre los libres no esclavistas. Comparar la comunidad real I en A con la comunidad real II en B, preguntando qué bienestar tendrían los individuos de I estando en B, es lógicamente semejante a (e incluso, más sencillo que) comparar la comunidad real en A con una comunidad hipotética en un estado hipotético B, preguntando qué bienestar tendrían los individuos reales en ese estado hipotético.

Cambios endógenos de los gustos

Algunos de los lectores pueden haber tenido la misma fantasía infantil que yo: una caminata en un paisaje en el cual se podía llegar de cualquier lugar a cualquier otro, sin subir ninguna colina, porque los contornos del paisaje cambiaban permanentemente, en función del camino que uno siguiera. Una fantasía modificada, pero todavía muy extravagante, podría ser la siguiente. De cualquier punto dado en un paisaje, uno podría moverse sin esfuerzo hacia cualquier punto situado en un nivel inferior, sin importar que hubiese colinas en el medio. También se puede suponer que mientras se va de un punto a otro punto de un nivel inferior, el paisaje cambia de manera tal que algunos de los puntos situados en el nivel superior ahora se tornan puntos del nivel inferior. Una cuestión intrigante sería preguntarse a *cuáles* puntos se puede llegar de esta manera desde un punto inicial dado a_1. Evidentemente, un punto b sería accesible si y sólo si hay una secuencia $a_1\, a_2 \ldots\, a_n b$, de manera que a_2 está cuesta abajo de a_1, a_3 cuesta abajo de a_2 ..., y b cuesta abajo de a_n. Esto, sin embargo, no es de gran ayuda a no ser que especifiquemos cómo, debido al paso de a_i a a_{i+1}, el conjunto de pasos cuesta abajo de a_{i+1} y el conjunto de pasos cuesta abajo de cualquier otro

punto también están cambiando. En términos modales tenemos que especificar cómo ocurren los cambios posiblemente posibles cuando realizamos uno de los posibles. En los sistemas modales del capítulo 1 el conjunto de mundos accesibles desde un mundo dado no está por supuesto, sujeto a cambio. En nuestra fantasía, sin embargo, no estamos limitados por esta concepción rígida de lo posible. Las posibilidades aparecen y desaparecen constantemente; algunas pueden perderse para siempre y otras, volver; algunas pueden ser adquiridas permanentemente, otras, de manera transitoria.

La *educación moral* puede ser entendida por medio de esa parábola. Los puntos de la colina son los estados respecto de los cuales se prefiere el presente estado. Los puntos del nivel más bajo representan una mejoría respecto del estado real. Mientras nos esforzamos para lograr mejorías a corto plazo nosotros mismos cambiamos, de modo que nuestra noción de qué es una mejoría puede no mantenerse igual. Puede resultar, por ejemplo, que de la situación a_1 vamos a a_2 subjetivamente vista como una mejoría; y de a_2 a a_3, igualmente vista como una mejoría; inclusive si en a_1 el paso a a_3 fue visto como un cambio hacia lo peor. Nosotros mismos o nuestros guardianes pueden explotar estos hechos de manera que nos lleven sin dolor ni esfuerzo hacia posiciones óptimas a largo plazo a través de una sucesión de pasos cortos, cada uno de los cuales es visto como una mejoría a corto plazo. Como nos decía Aristóteles,[36] nuestras acciones surgen de nuestro carácter y nuestro carácter está formado por nuestras acciones. En palabras del economista: "La elección depende del gusto y el gusto, de elecciones pasadas."[37] Esto muestra una promesa de mejoría y un peligro de inconstancia; una herramienta para la autodidáctica y, asimismo, para la manipulación. La característica principal de este método de aproximación a la acción humana es que une dos aspectos, con frecuencia pensados como incompatibles: la intencionalidad y la causalidad. La elección de acuerdo con las preferencias firmes es una parte importante de lo que nosotros entendemos por intencionalidad, mientras que las leyes que gobiernan los cambios endógenos de las preferencias son puramente causales, operando "a espaldas" del actor. El actor puede, por consiguiente, intentar anticipar el impacto del comportamiento presente sobre las preferencias futuras y hacer su elección presente a la luz de su evaluación de esas consecuencias; en verdad, esto es lo que Aristóteles nos dice que hagamos.[38] Esto, sin embargo, hace que surja el mismo problema más adelante porque, incluso si de esa manera podemos escoger nuestras preferencias, sigue estando la posibilidad de un cambio endógeno de nuestras metapreferencias, y así sucesivamente. Volveré a este problema en el apéndice 2 del capítulo 5.

El análisis económico del cambio endógeno del gusto se remonta a 1944, cuando Trygve Haavelmo abrió perspectivas totalmente nuevas, con breves y sagaces observaciones en su disertación doctoral.[39] Desarrollos posteriores importantes están ligados a los nombres de Nicholas Georgescu-Roegen, C.C. von Weizsäcker y Robert Pollak.[40] Aquí bosquejaré algunas de las ideas del seminal trabajo de Von Weizsäcker. Por simplicidad supone que, (i) las preferencias actuales dependen sólo del consumo en el

período inmediatamente anterior y (ii) el ingreso del consumidor debe distribuirse solamente sobre dos bienes. Recientemente se ha demostrado[41] que la segunda presunción es decisiva para los resultados obtenidos por Von Weiszäcker, así que no se generalizan fácilmente en el caso de *n*-mercancías. No tomaré este problema en cuenta pues pienso que hay cuestiones filosóficas importantes a discutir, incluso en aquellos casos más especiales en los que los resultados *son* válidos. La primera presunción es más bien perturbadora, ya que excluye la posibilidad de una formación del carácter acumulativo (en verdad, partes del propio desarrollo de Von Weiszäcker entran en contradicción flagrante con su presunción).[42] Sin embargo, está claro que la introducción de toda la historia del consumo como variable complicaría el análisis haciéndolo virtualmente inmanejable y que, en el estado actual de la investigación, la estrategia escogida por Von Weiszäcker es probablemente la correcta.

Los economistas, cuando estudian la elección del consumidor, utilizan tres conceptos de manera más o menos intercambiable: preferencias, funciones de utilidad y funciones de demanda. Dadas diversas condiciones, que suponemos cumplidas, dos de estos conceptos cualesquiera pueden ser deducidos o formados a partir del tercero.[43] Permítasenos suponer que el consumidor en el momento t tiene una preferencia por la estructura P_t, representada por una función de utilidad u_t de la cual podemos derivar una función de demanda f_t que a cada conjunto de precios e ingresos correlaciona un cierto vector de mercancía (de dos componentes) escogido por el consumidor: $\mathbf{q}_t=f_t(\mathbf{p},y)$. Suponemos ahora que la demanda actual es una función de demanda y consumo en el período inmediatamente anterior: $f_t=F(f_{t-1}, \mathbf{q}_{t-1})$ donde F es una función que representa los rasgos de carácter constantes que subyacen al cambio de preferencias. Más específicamente, presumimos que F es tal que existe un g tal que $\mathbf{q}_t=g(\mathbf{p}, y_1\, \mathbf{q}_{t-1})$. Esta última fórmula es sólo utilizada por Von Weiszäcker. Ahora, las aplicaciones sucesivas de la función g generan una secuencia $\mathbf{q}_i$ de haces de mercancías, que bajo determinadas condiciones (cuya interpretación económica no es evidente) se demuestra que converge en un punto $\mathbf{q}=h(\mathbf{p},y)$, que es independiente del haz inicial. (Este resultado demuestra que el peligro de inconstancia y cambios interminables de preferencias puede evitarse por lo menos en algunos casos.) La relación h puede ser formalmente interpretada como una función de demanda ya que correlaciona a cada conjunto de precios e ingresos con un cierto haz de mercancías. A esta función de demanda (dadas ciertas suposiciones cruciales y controvertibles) corresponde una función de utilidad U "como si" en el sentido del que podemos deducir el paquete $\mathbf{q}=h(\mathbf{p},y)$ si presumimos que el consumidor está maximizando U, dadas $\mathbf{p}$ e y. Alternativamente, podemos construir la "estructura de preferencias a largo plazo" P, con la propiedad de que $\mathbf{x}P\mathbf{y}$, si y sólo si $U(\mathbf{x}) > U(\mathbf{y})$.

Volvamos ahora a las preferencias "instantáneas" P_t. En verdad, no hay necesidad de indexarla por el *tiempo*, porque la única diferencia pertinente entre P_{t1} y P_{t2} es el consumo en los períodos inmediatamente precedentes t_1-1 y t_2-1; si el consumo en estos períodos fue el mismo,

entonces las preferencias instantáneas en los períodos siguientes serán también las mismas. Podemos por lo tanto indexar de manera más útil las preferencias por el consumo en el período precedente, escribiendo "$\mathbf{x}P(\mathbf{z})\mathbf{y}$" para "**x** es preferida a **y**, dado que el consumo en el período inmediatamente precedente fue **z**". Von Weiszäcker demuestra ahora el siguiente teorema, válido bajo ciertas condiciones, que puede ser verbalmente expresado como inercia o conservadorismo del consumidor:

Dados dos vectores de mercancías **x** *e* **y**, *entonces* **x** *P***y** *si y sólo si existe una secuencia de vectores* $\mathbf{r}_1 ... \mathbf{r}_n = \mathbf{x}$ tal que $\mathbf{r}_1 P(\mathbf{y})\mathbf{y}$, $\mathbf{r}_2 P(\mathbf{r}_1)\mathbf{r}_1$, $\mathbf{r}_3 P(\mathbf{r}_2)\mathbf{r}_2 ... \mathbf{r}_n P(\mathbf{r}_{n-1})\mathbf{r}_{n-1}$.

El teorema dice, en otras palabras, que **x** es mejor que **y** de acuerdo con mis preferencias a largo plazo si y sólo si existe un camino de **y** a **x** tal que cada paso en el camino es una mejoría de acuerdo con las preferencias instantáneas al inicio de ese paso. Dos observaciones deben hacerse en este momento. Primero, puede bien ser el caso de que $\mathbf{y}P(\mathbf{y})\mathbf{x}$, lo cual implica un conflicto entre las preferencias a corto y a largo plazo. El teorema, entonces, afirma que tal conflicto puede ser superado, utilizando estrategias indirectas: es la fantasía infantil (en su versión modificada) que se torna verdadera. Segundo, puede darse el caso de que, aunque ambas **x** e **y** estén dentro del conjunto factible para el consumidor, algunos vectores **r** pueden ser imposibles de obtener a los precios e ingresos dados. En este caso, el consumidor no puede educarse a sí mismo para realizar sus preferencias a largo plazo o, más bien, su educación tiene que ser una batalla cuesta arriba en vez de una caminata cuesta abajo. El gobierno o alguna otra autoridad puede, sin embargo, subsidiar su mejoramiento, haciendo que los espacios intermedios estén dentro de su conjunto factible. Von Weiszäcker da como ejemplo el problema de la movilidad regional, donde el gobierno puede explotar su conocimiento acerca del cambio endógeno de preferencias, para subsidiar una transferencia indolora del campo a la ciudad.

Parecería que, de acuerdo con Von Weiszäcker, esta clase de manipulación a través de subsidios es siempre éticamente aceptable; pareciera que él también piensa que el individuo no tendría base para oponerse a la manipulación. Imagínese, sin embargo, la siguiente situación. [44] Yo soy un no fumador pues no me gusta el tabaco barato, que está dentro de las posibilidades de mi presupuesto. Sin embargo, si alguien me subsidiara para comprar tabaco mejor, mi ansiedad por la nicotina me llevaría a aceptar gustosamente, después de un tiempo, el tabaco barato que antes rechazaba. ¿Sería éticamente justificable para alguien, un productor de tabaco, digamos, ofrecerme tales subsidios y no tendría yo que tener escrúpulos para aceptarlos? Ambas preguntas tendrían que ser presumiblemente contestadas por la negativa (si no está convencido de esto, sustituya el tabaco por heroína) lo cual muestra que mis preferencias a "largo plazo" no son mis preferencias "reales". Seducir a alguien y llevarlo a hacer algo que va contra sus intereses reales no es más justificable que coercionarlo.

(Entre el caso puro de seducción, donde cada paso se persigue como una mejoría, y el caso puro de coerción, donde el comportamiento sólo puede ser mantenido con una permanente violencia, hay un caso intermedio interesante de acción coercitiva que, rápidamente, se vuelve formadora-de hábitos y, por lo tanto, se torna superflua. Esto, quizá, se correspondería con lo que los moralistas y novelistas han tenido en mente al discutir la seducción, como en *Las Relaciones Peligrosas*.)

Los casos de razonamiento posibilístico discutidos en este capítulo no tienen, virtualmente, nada en común, excepto que de una manera u otra las potencialidades puras desempeñan un papel importante. En el contexto del presente trabajo, en conjunto, estos casos cumplen una doble función. Primero, sirven para recordar al lector que las ciencias sociales ya "habitan en la Posibilidad": que un número de teorías muy conocidas concierne a la relación entre lo que los individuos y los grupos hacen realmente y lo que pueden hacer (o haber hecho). Segundo, algunos de los ejemplos específicos serán utilizados en discusiones posteriores. En particular, el análisis del ingreso nacional real que nos interesará nuevamente en el capítulo 6, mientras que el problema del cambio endógeno de los gustos se toca tanto en el capítulo 4 como en el apéndice 2 del capítulo 5.

APÉNDICE DEL CAPÍTULO 2

Un resultado de imposibilidad para gramáticas independientes del contexto

En la sección sobre las gramáticas de estados finitos, probamos que son inadecuadas en principio para la generación de lenguajes naturales. Un argumento similar puede utilizarse para probar la inadecuación de lo que Chomsky [45] llama gramáticas Tipo 2 o gramáticas independientes del contexto. De acuerdo con mis conocimientos, esta independencia del contexto no ha sido nunca probada en la literatura, a pesar de afirmaciones en contrario.[46] Como una introducción a la prueba debemos explicar la diferencia entre gramáticas de estados finitos y gramáticas independientes del contexto. Supongamos que una frase $w_1 w_2 w_3 w_4 w_5$ ha sido generada por una gramáticas de estados finitos, pasando a través de la siguiente secuencia de estados: s_1-s_2-s_3-s_4-s_5-s_1. Esto también puede expresarse diciendo que la oración ha sido generada por la sucesiva aplicación de las siguientes *reglas de reescritura*:

$$s_1 \rightarrow w_1 s_2,$$
$$s_2 \rightarrow w_2 s_3,$$
$$s_3 \rightarrow w_3 s_4$$
$$s_4 \rightarrow w_4 s_5,$$
$$s_5 \rightarrow w_5.$$

Estas reglas tienen la característica formal de que a la izquierda vemos un solo símbolo abstracto, que luego se transforma bien sea en una palabra o en una palabra seguida de un símbolo abstracto. Esto, en verdad, provee una caracterización independiente de la clase de gramáticas de estados finitos. Una gramática independiente del contexto permite reglas más complejas: a la izquierda puede verse un solo símbolo abstracto, pero éste puede transformarse en una concatenación de símbolos y palabras abstractas (más complejas aun son las gramáticas dependientes del contexto, que permiten reglas de tipo $ABC \rightarrow ADC$, donde B es un solo símbolo abstracto y A,C,D puede ser cualquier concatenación de símbolos y palabras abstractas. No hay prueba de la inadecuación de dichas gramáticas.) Para estas gramáticas independientes del contexto, Bar-Hillel y cols.[47] han probado el siguiente teorema:

Sea A una gramática independiente del contexto y L (A) el conjunto de oraciones generadas por A. Entonces, existe un número p tal que para cualquier z en L (A) más larga que p existe x,u,w,v, y tal que

(i) $z=xuwvy$

(ii) $xu^m wv^m y \varepsilon L(A)$ *para toda m;*

(iii) u y v no son ambas vacías.

Más brevemente, el teorema afirma que en un lenguaje generado por una gramática independiente del contexto cualquier oración suficientemente larga puede ser extendida por una repetición indefinida de *dos* subpartes, sin violación de la gramaticalidad, tal como las gramáticas de estados finitos en las que cualquier oración suficientemente larga puede ser extendida por la repetición indefinida de una subparte. Para probar la inadecuación de las gramáticas independiente del contexto para los lenguajes naturales, tenemos que traer un contraejemplo de una oración gramatical que no pueda ser prolongada de esta manera. Los contraejemplos de las formas "Si ^{k} sopla viento (entonces llueve)^k" no sirven para este fin, porque estas oraciones *pueden* ser extendidas indefinidamente por la repetición de los dos bloques exponenciales. Ahora, veamos las siguientes oraciones:

B_1: los primeros dos millones de números en la expansión decimal de π son $a_1\ a_2 \ldots\ a_{2000000}$.

B_2: los primeros dos millones de millones de números en la expansión decimal de π son $a_1\ a_2 \ldots\ a_{20000000000000}$.

.
.
.

B_k: los primeros dos (millones)^k de números en la expansión decimal de π son $a_1\ a_2 \ldots\ a_{2.10^{6k}}$.

.
.
.

Dada una gramática independiente del contexto y el número p del teorema anterior es sencillo encontrar un k tal que B_k es más larga que p. Tal B_k tiene que permitir una doble extensión, como se afirma en el teorema. No es difícil ver que estas extensiones tienen que ocurrir dentro de los bloques "millón^k", y "$a_1....a_{2.10^{6k}}$", si el resultado no ha de ser obviamente no gramatical. Supongamos, entonces, que u en el teorema tiene la forma "millón^q" y v, la forma "$a_r...a_t$". Eligiendo $m = 2$ en el teorema, entonces afirma que la siguiente oración debe ser gramatical si suponemos que la gramática independiente del contexto en cuestión es adecuada para el inglés:

C: los primeros dos millones $^{k+q}$ de números en la expansión decimal de π son $a_1...a_t a_r...a_t...a_{2.10^{6k}}$

Demostraré que *C* no es una oración gramatical en inglés. Si *C* es gramatical, entonces el número "dos millones^{k+q}" tiene que ser igual que el número "$2.10^{6k} + t - r + 1$", es decir, igual que el número de números en la expansión decimal. Obsérvese que ésta es una exigencia no de las matemáticas sino de la lingüística, igual que la falta de gramaticalidad de la oración,

D: Los dos animales más grandes del zoológico son un ratón,

es una cuestión de la lingüística y no de las matemáticas (o de la zoología), pero como $q \geq 1$ y $t - r + 1 \leq 2.10^{6k}$, tenemos

$$2.10^{6k} + t - r + 1 \leq 4.10^{6k} < 2.10^{6k+1} < 2.10^{6(k+q)},$$

mientras que la gramaticalidad de *C* requeriría $2.10^{6k} + t - r + 1 = 2.10^{6(k+q)}$.

El resultado que acabamos de comprobar, además del interés intrínseco de la imposibilidad, tiene la ventaja de resaltar (incluso más claramente que el contraejemplo de las gramáticas de estados finitos) la naturaleza puramente posibilística del razonamiento lingüístico. Tenemos nociones claramente intuitivas de la gramaticalidad de las oraciones con probabilidad cero de ser pronunciadas, e incluso debe ser rechazada una gramática que podría generar exactamente un conjunto de oraciones con una probabilidad no cero de ser pronunciadas, si fracasa en generar una sola oración gramatical como B_k.

Notas

1. Véanse las alertas de Duncan (1966).
2. Véase Fararo (1973 págs. 142 y sigs.).
3. Este era el punto de vista de Sombart, citado por Lipset y Bendix (1959, pág. 12).
4. Véase Fararo (1973, págs. 147 y sigs. y 277 y sigs.). Si *D* es la matriz adyacente, la matriz de alcanzabilidad es $D + D^2 ... D^{n-1} + D^n$.
5. Confróntese Mc Farland (1970) para una elegante observación.
6. Véase Coleman (1973a), Boudon, (1973) y Fararo (1973) para exposiciones diversas.

7. Véase Feller (1968, pág:122). Utiliza el ejemplo de que los automovilistas que han tenido un accidente tienen más posibilidades de tener más accidentes, lo cual puede ser interpretado en términos de "conductores naturalmente malos" o en términos de conductores inicialmente indiferenciables que son selectivamente impactados por su primer accidente hacia conducir mal. Otro ejemplo puede ser el desempleo: dado el conocimiento empírico de que la probabilidad de encontrar un trabajo disminuye con la duración del desempleo (Cripps y Tarling, 1974) esto puede ser comprendido, bien sea como resultado de que algunos individuos son "por naturaleza" menos emprendedores y capaces que otros, o de la hipótesis de que la gente ha cambiado (para empeorar) por el desempleo, aunque inicialmente hubiera perdido el empleo por un mecanismo casual. Esta creación irreversible de diferencias a través de opciones azarosas en una población inicialmente homogénea puede ser un mecanismo importante para explicar el cambio social.

8. Confróntese McFarland (1970) para una observación sagaz de los procesos del método Markov como analogía estocástica a los procesos deterministas.

9. Confróntese Elster (1976a) para una discusión más completa.

10. Chomsky y Miller (1963, pág.429).

11. Chomsky y Miller (1963, pág.429).

12. Cf. Hockett (1955, pág.10), citado por Chomsky (1957, pág. 17): Debemos reemplazar "la probabilidad cero y todas las probabilidades extremadamente bajas por *imposible* y todas las probabilidades mayores por *posible*". Esto significaría que las posibilidades son definidas como *probabilidades truncas.*

13. La siguiente prueba se toma de Rabin y Scott (1959).

14. Lévi-Strauss (1958, págs.311 y sigs.).

15. Para dos lecturas divergentes véase Nutini (1970) y Barnes (1971), de las cuales, la última está cercana a mi lectura de Lévi-Strauss.

16. Lévi-Strauss (1973, págs. 351 y sigs.).

17. "Las ciencias sociales están en una relación de cliente con respecto a las ciencias exactas y naturales, mientras que las ciencias humanas aspiran a una relación de discípulo" (Lévi-Strauss, 1973, pág. 362).

18. "Las ciencias humanas son aquellas que se sitúan a sí mismas por fuera de cada sociedad particular; o bien porque adoptan la perspectiva de una sociedad en general (*une société quelconque*) o de un individuo en general, en una sociedad en general (*un individu quelconque au sein de n'importe quelle société*), o bien porque están tratando de asirse a una realidad humana inmanente y, por lo tanto, situándose a sí mismas en una perspectiva que es previa, tanto al individuo como a la sociedad" (Lévi-Strauss, 1973, pág.360).

19. "De esta manera, uno podría obtener una fuente capaz de generar una tabla periódica de estructuras lingüísticas, similar a la tabla periódica de elementos que la química moderna debe a Mendeleieff. Sólo tendríamos entonces que ubicar los lenguajes ya estudiados en esta tabla, para observar su posición y relación con otros lenguajes que son teóricamente poco conocidos, y para descubrir la ubicación de otros lenguajes, pasados, futuros o simplemente posibles" (Lévi-Strauss, 1958, pág.66).

20. Lévi-Strauss (1967, pág.98).

21. Jakobson y Halle (1963, págs. 145 y sigs.).

22. Lévi-Strauss (1958, pág.99).

23. Lévi-Strauss (1958, pág.233); confróntese también Lévi-Strauss (1964–71).

24. Lévi-Strauss (1958, capítulo II).

25. Lévi-Strauss, (1973, pág.107). Para una respuesta diferente y más bien incompatible con esta objeción véase Needham (1962, pág. 35).

26. Véase Kemeny y cols. (1966, págs.424 y sigs.) y Fararo (1973, capítulo 18).

27. Véase también Bourdieu y cols. (1968, pág. 80): "El tipo ideal debería ser tratado no como un objeto autónomo (*non en lui–même et pour lui–même*) sino como un miembro de un grupo en transformación y en su relación con todos los otros miembros de la familia de la cual es un miembro privilegiado."

28. Para este ejemplo véase Balibar (1966, pág.211).

29. Samuelson (1950). Citado de la reproducción de Arrow y Scitovsky (1969, pág. 418).

30. Para un ejemplo numérico que exhiba la posibilidad de esta situación véase Quirk y Saposnik (1968, pág.121).

31. Para un ejemplo numérico véase Quirk y Saposnik (1968, pág. 122).

32. Confróntese Graaff, (1971, págs. 77 y sigs.), para una exposición sobre este tema. La dificultad es más o menos la siguiente. Para que funcione una solución al bulto, la gente no debe ser inducida a ocultar sus reales capacidades, lo que presupone una solidaridad que hubiera impedido que surgiera en primer lugar el problema para el cual el impuesto es una solución. Confróntese también Sen (1973, capítulo 4).

33. Arrow y Scitovsky (1969, pág. 385).

34. Sen (1976).

35. Sen (1976) extiende el análisis al caso de las comunidades de diferentes tamaños.

36. *La ética nicomaquea*, 1113–1114. Siguiendo a Kenny (1963), podemos explicar la teoría de la educación moral de Aristóteles, en los siguientes términos. El carácter es algo que se adquiere a través de la acción. Aristóteles se refiere a este carácter como *hexis*, algo que se *posee*. Las acciones que establecen esta *hexis* caben bajo el rótulo de *kinesis* o *poiesis*, por lo menos (me parece) cuando se hacen para establecer esta *hexis*. Las acciones que parten de la *hexis* entran en la categoría de energeia o praxis; no son un medio hacia un fin sino que poseen el fin en sí mismas. El hombre bueno *se hace* bueno haciendo acciones buenas (*kinesis*); cuando es bueno, estas acciones (*energeia*) fluyen sin ninguna restricción desde su ser. Hacerse bueno es una batalla cuesta arriba, ser bueno es como el agua que busca su curso natural cuesta abajo. Para un modelo simple de relación entre *kinesis, hexis y energeia*, podemos tomar la relación entre aprender francés, saber francés y hablar francés.

37. Gorman (1967, pág. 218).

38. "Por lo tanto, sólo una persona totalmente insensata puede ignorar que nuestro carácter es el resultado de la conducta; pero si un hombre actúa conscientemente, de manera tal que su acción es propia de un hombre injusto, es necesario decir que es voluntariamente injusto."(Aristóteles, *loc. cit.*). Hay una seria ambigüedad aquí, pues incluso si conozco (de manera general) *que* mis acciones presentes van a tener un impacto sobre mi carácter futuro puedo no estar capacitado para decir *cómo* van a cambiar mis preferencias. En el capítulo 4 argumento que ésta es una característica de la que no puede escapar la *condition humaine*; que aunque con frecuencia sepamos que nuestras acciones van a tener consecuencias importantes, no estamos en condiciones de decir cuáles consecuencias; que aunque sepamos que en un tiempo posterior vamos a arrepentirnos de nuestras acciones presentes, no sabemos cuáles acciones; que aunque podamos afirmar en el prefacio de un libro que éste contiene errores no estamos en posibilidad de señalarlos. La imposibilidad de tener un conocimiento específico acerca del impacto del comportamiento presente sobre los rasgos futuros del carácter no parte de la incertidumbre acerca del medio ambiente ni de las dificultades prácticas para establecer las conexiones causales pertinentes. Estamos tratando con una imposibilidad absoluta, porque la articulación causal entre nuestras acciones y los rasgos de primer orden de nuestro carácter es en sí misma un rasgo de segundo orden del carácter, que puede ser modificado por las acciones, y así sucesivamente en una secuencia siempre ascendente. Es imposible, en principio, tener un conocimiento específico de los niveles de esta jerarquía infinita, aunque es posible tener conocimiento general acerca de la existencia de una jerarquía y un conocimiento concreto acerca de alguno de sus niveles. Confróntese también con el apéndice 2 del capítulo 5 y Elster (1976d).

39. Haavelmo (1944, págs.17 y sigs.).

40. Georgescu-Roegen (1950); Von Weiszäcker (1971a); Pollak (1976a).

41. Por Pollak (1976a).

42. En la sección 5 de su trabajo, Von Weiszäcker (1971a) pregunta por qué "el comportamiento de la demanda a largo plazo debería ser distinto del comportamiento de la demanda a corto plazo", y responde que el cambio se debe a que el consumidor ha pasado por un proceso de aprendizaje. Puede empezar siendo más bien ignorante de las diversas posibilidades de consumo y cambiar fácilmente su modo de pensar en el sentido de que las evaluaciones *ex post* van a diferir de las *ex ante*. Después de algún tiempo, sin embargo, habrá recorrido grandes regiones del mercado y formado preferencias estables, que reflejan su

experiencia. Esta es, en verdad, una idea muy importante, que podría ser incorporada a la teoría del cambio social endógeno (confróntese Elster, 1976a, para un bosquejo de cómo esto puede hacerse). Es incompatible, sin embargo, con la muy corta memoria postulada en todo el resto del trabajo. Si el consumidor sigue una secuencia de vectores de consumo de $\mathbf{q}_1 \mathbf{q}_2 \mathbf{q}_3 \mathbf{q}_4 \mathbf{q}_5 \mathbf{q}_1 \mathbf{q}_2$, el estado de la mente del consumidor cuando llega a $\mathbf{q}_2$ la segunda vez no refleja la experiencia de $\mathbf{q}_3 \mathbf{q}_4 \mathbf{q}_5$ si postulamos $\mathbf{q}_1 = g(\mathbf{p}, y, \mathbf{q}_{t-1})$.

43. Para una exposición véase, por ejemplo, Georgescu-Roegen (1968)

44. Este contraejemplo fue independientemente dado por Pollak (1976a, págs. 294-295). El-Safty (1976, pág.300) ofrece el siguiente ejemplo: "A Johnny se le preguntó por qué no comía espinacas. Contestó «No me gustan las espinacas». A esto, un adulto respondió: «Pruébalas, te gustarán». «Ya sé», contestó Johnny, «pero no quiero que me gusten» ". Para que este ejemplo sea convincente, tenemos que asumir que lo que Johnny verdaderamente no quiere son los efectos colaterales de comer espinacas, tales como adquirir la fuerza de Popeye, que lo haría susceptible de realizar más tareas domésticas u otras cosas.

45. Chomsky (1959).

46. Postal (1964, especialmente págs. 145 y sigs.) afirma haber probado que las gramáticas independientes del contexto no pueden generar el lenguaje Mohawk. Se refiere a la prueba de Chomsky (1959) del siguiente teorema:

Sea K un alfabeto con al menos dos símbolos, y A cualquier conjunto finito de palabras sobre K. Entonces, el conjunto diagonal $D_A{}^2$ (es decir, el conjunto de todas las cadenas XX, donde X es un miembro de A) no puede ser generado por una gramática independiente del contexto.

Postal afirma en un pie de página (nota 27) que este resultado puede generalizarse al siguiente teorema:

Sea K un alfabeto por lo menos con dos símbolos y A un conjunto infinito de palabras sobre K. Entonces, el conjunto diagonal $D_A{}^2$ no puede ser generado por una gramática independiente del contexto.

Esto, sin embargo, es erróneo. Para un contraejemplo tómese $K = (a, b)$ y $A = ((ab), (ab)^2 ... (ab)n...)$. Entonces, $D_A{}^2 = ((abab)... (abab)^n...)$ la cual es generada po $S \rightarrow abab$, $S. \rightarrow ababS$, una gramática independiente del contexto. Para que su argumento tenga validez, Postal requiere la versión generalizada del teorema o, por lo menos, tendría que demostrar que el teorema es válido para el conjunto particular *A* en el Mohawk escogido para la discusión (le debo la sustancia de esta nota a la sugerencia de Aanund Hylland).

47. Bar-Hillel y cols. (1964).

3

Posibilidad política

La lógica modal es la ciencia de lo posible, la política es el arte de lo posible, y nunca ambas se han encontrado. En este capítulo intento una síntesis parcial de los dos campos por medio de una interpretación política de las nociones básicas de la lógica modal. No creo que este ejercicio pueda dar una visión nueva y sustancial de la naturaleza de la política ni, tampoco, que podamos sacar ningún conejo empírico del sombrero de la lógica. La justificación de lo que voy a hacer es, primero, que la escasa formalización que intento puede conducir a explicar la conexión entre nociones que, a primera vista, no parecen relacionadas y, segundo, que los conocimientos modales pueden llevarnos a formular algunas preguntas nuevas que, de otro modo, no hubieran surgido naturalmente. La lógica modal, aquí o en cualquier otro lugar, es a lo sumo una máquina que genera preguntas y no un atajo hacia la obtención de respuestas. Además de ser pertinente para los científicos políticos, pienso, no sin cierta timidez, que esta utilización de la teoría de la lógica puede ser de interés para el lógico en búsqueda de material ilustrativo e incluso, quizás, un incentivo al desarrollo conceptual.

El postulado básico de la aproximación es que en un estado dado de la sociedad existe un conjunto de estados políticamente posibles con respecto a ese estado. Utilizaré la abreviatura 'PP c.r.a' para la relación de accesibilidad política. En lugar de hablar acerca de los estados mismos, quizás es más apropiado describirlos de algún modo determinado, por ejemplo, por medio de un conjunto de variables. Para un estado dado —como una entidad ontológica— presumiblemente hay a lo sumo un estado que es PP c.r.a ese estado (más adelante veremos por qué no sería correcto decir exactamente un estado), pero para el estado descrito de una determinada manera el conjunto de estados que son PP c.r.a puede tener más de un miembro. Esta subdeterminación del futuro surge de la subespecificación del presente, que es un rasgo inevitable de una descripción finita. Si el lector no olvida esta diferenciación entre estado y descripción del estado, no se sorprenderá por estas afirmaciones *prima facie* extrañas como que el conjunto de posibilidades políticas en un estado dado puede cambiar

incluso cuando el estado mismo permanece igual. De manera similar, la histéresis o acción temporal a distancia, que es imposible a nivel ontológico, puede, sin embargo, figurar en el nivel epistemológico como una característica de nuestras teorías acerca de determinados procesos.[1]

El enfoque que se presenta está centrado en el actor, en el sentido de que sólo me refiero a los estados que pueden ser *causados* deliberadamente, mientras excluyo los que *ocurren* causalmente. Cualquier estado dado tiene que estar seguido por algún otro estado (o por sí mismo), de manera que si la posibilidad política se definiera en sentido puramente causal, el conjunto de estados que son PP c.r.a un estado dado nunca estaría vacío. Argumentaré, sin embargo, que puede haber estados con la característica peculiar de que la acción política intencional está condenada a fracasar, o sea, que no hay posibilidad, en el sentido intencional del término.

Ahora, podemos usar con un propósito útil una analogía con el poder. En general, no decimos que el poder es sólo una producción de resultados, un comportamiento cuya consecuencia es hacer que las personas actúen de una manera diferente a como lo hubieran hecho.[2] Más bien, queremos decir con Bertrand Russell[3] que el poder es la producción de resultados *intencionados* o, quizá, la capacidad para producir dichos resultados. En esta definición son esenciales los aspectos causales y los intencionales. El motivo por el que el aspecto causal es esencial está bien resaltado por Alvin Goldman: "Decir que S es poderoso no es decir que usualmente obtiene lo que de hecho quiere, sino que podría obtener cualquier cosa que quisiera, sin importar lo que fuera."[4] Sin la última cláusula de esta definición, tendríamos que adscribirle poder extremo al camaleón, es decir, a la persona que siempre cambia sus preferencias para estar del lado ganador. Estoy de acuerdo con Goldman en que ignorar este aspecto nos conduce, contra la intuición, a atribuirle poder a personas que no lo tienen, aunque en el capítulo 6 argumentaré que el aspecto contrafáctico del poder crea problemas que, por el momento, me parecen insolubles. Ahora bien, en el caso del camaleón, el aspecto intencional está presente y el causal, ausente. A la inversa, una persona puede estar principalmente ubicada sobre lo causal, en el sentido de que lo que haga sea decisivo para el resultado de la situación, pero ser incapaz de obtener el resultado deseado. Un ejemplo, en pequeño, sería el capitán de un barco perdido en el mar sin posibilidad de navegar; otro ejemplo mayor sería un gobierno que contemple la adopción de diversas medidas contra la inflación. Si alguien tiene poder para contener la inflación, es el gobierno; está en el lugar del conductor. Sin embargo, si no conoce las relaciones fines-medios que se requieren, la cláusula "si alguien..." se satisface de manera vacua; falta el elemento intencional, aunque esté presente el causal.

Es posible avanzar un paso más y afirmar que, incluso la producción causal de resultados intencionados no es siempre un ejercicio de poder, por ejemplo, si los resultados surgen de un mecanismo diferente del que el actor tenía en mente. De la misma manera que el conocimiento no es sólo

creencia verdadera sino creencia verdadera *justificada*, el poder no es sólo la producción causal de resultados intencionados sino la producción causal *inteligente* de resultados intencionados. El gobierno puede pretender contener la inflación con medidas $X_1...X_n$ y al mismo tiempo intentar alcanzar otros fines por medio de medidas $Y_1...Y_m$. Si la inflación es, en verdad, reducida porque *o bien* $(X_1...X_nY_1...Y_m)$ es una causa suficiente mientras $(X_1...X_n)$ no lo es *o* porque $(Y_1...Y_m)$ por sí sola es causa suficiente, mientras $(X_1...X_n)$ no lo es, entonces deberíamos dudar en decir que el gobierno ha ejercido su poder para frenar la inflación. (Si el gobierno piensa que esto es lo que ha hecho y actúa de igual manera cuando de nuevo haya inflación, estaríamos frente a un caso de "aprendizaje supersticioso").[5]

La noción de posibilidad política depende, también, de la producción intencional e inteligente de estados deseados, por razones análogas a las del poder. Además, no es orientada por el actor, en el sentido de restringirse a lo que es visto *ex ante* como PP por los actores mismos. Este es un punto fundamental, que debe estudiarse extensamente. Permítasenos observar primero que, en verdad, hay en la vida política un aspecto de luchas rutinarias por el poder, dentro de límites aceptados y comprendidos por todos. Dentro de esos límites están los resultados posibles, por los cuales es válido entrar en la pelea; por fuera, están las imposibilidades políticas, que nadie en sus cabales querría convertir en su problema. En términos de Spinoza, este comportamiento rutinario es *política politizada*, opuesta a la *politización de la política* en la que se crean las rutinas. (Diremos más acerca de esta diferenciación en el capítulo 5.) En la primera, la lucha ocurre, de común acuerdo, dentro de las fronteras de lo posible; en la última, la lucha es para redefinir estas fronteras. Hay una diferenciación similar en la lingüística entre "creatividad gobernada por normas" y "para el cambio de norma";[6] ambas se manejan con posibilidades, pero sólo la primera lo hace con posibilidades explicables por la ciencia. Sartre, cuando escribió que la conciencia *'est été'*,[7] indiscutiblemente violó la gramática francesa, pero al hacerlo también creó una nueva posibilidad gramatical, que se sumó al repertorio del lenguaje. La gramática política puede ser violada y reconstruida de manera similar por la invención de nuevas instituciones políticas, tales como la empresa multinacional; por formas nuevas, tales como la guerra de guerrillas; por conceptos políticos nuevos, como el impuesto negativo a los ingresos.

Mientras la teoría de la posibilidad política esté centrada en el actor en el sentido de lo que ahora está en discusión, sólo podría ser una teoría de política rutinaria. O sea, confinarse al conjunto de estados que todos reconocen *ex ante* como PP c.r. al estado dado, excluyendo el proceso creativo donde lo posible sea *conquistado* más bien que realizado. Un ejemplo de la teoría revolucionaria marxista puede ayudar a clarificar este punto. En esa teoría se le atribuye un papel importante al "análisis de la situación", es decir, a la evaluación del potencial revolucionario en una situación dada. El propósito de ese análisis no es sólo reflejar o describir la

situación, para utilización de los dirigentes revolucionarios, sino ser un arma para la transformación de la situación misma, por medio del impacto sobre los actores políticos, cuando el análisis se publica. Diciendo que la situación *es* revolucionaria se puede hacer que ésta se *transforme* en revolucionaria, por medio de una cierta profecía que se cumple a sí misma. En general, sabemos que diciéndolo no lo hacemos y que sólo excepcionalmente será posible alcanzar metas políticas por medio del pensamiento de lo deseado.[8] Incluso, cuando *es* posible ubicar la revolución en el mapa diciendo que está allí, esta posibilidad no puede ser delineada con anterioridad. Permítasenos suponer que fuera imposible especificar las características que permiten al dirigente transformarla de una situación en la cual la revolución por consenso universal es imposible a otra situación en la cual la revolución se reconoce como una posibilidad real, simplemente afirmando esta posibilidad. Podemos llamarla situación prerrevolucionaria, significando algún tipo de descontento latente o parcialmente cristalizado que necesita sólo la palabra correcta en el momento correcto para emerger en forma reconocible. En ese caso, los dirigentes buscarían naturalmente las características de una situación que les permitiera hacer que la situación fuera prerrevolucionaria simplemente afirmando que lo es, y así en un retroceso infinito, que deja ver el carácter esquivo de las posibilidades políticas *ex ante*. Entre lo que es posible sin ninguna ambigüedad y lo que es imposible sin ninguna ambigüedad, hay un limbo en el que sólo la *acción* puede decidir. Por otra parte, el científico social puede establecer con precisión la línea divisoria, con la condición de permanecer separado del evento que estudia.[9]

Sobre el principio "deber implica poder", parecería natural requerir a un estado dado que es PP c.r.a un estado dado que sea económicamente factible en relación con ese estado. En el capítulo 2 me referí a la noción de Paul Samuelson acerca de la distribución de ingresos políticamente factible, donde se toma como real que éstos forman un subconjunto de estados económicamente factibles. Aquí también impondré esa exigencia, pero podemos detenernos para señalar una conceptualización alternativa, en la que los estados que son PP c.r.a un estado dado son los estados que pueden ser *aceptados* como metas políticas incluso si no pueden realizarse. Hasta hace poco, ningún partido político de Europa occidental podía afirmar abiertamente que el pleno empleo, el crecimiento económico y los precios estables eran fines que no podían realizarse simultáneamente en el marco de la moderna economía capitalista. Por lo tanto, todos los partidos políticos tenían que declarar las imposibilidades económicas como posibilidades políticas; evidentemente eran *necesidades* políticas (en el sentido que ahora discutimos) ya que eran comunes a todos los partidos. (El análisis de Kalecki-Lindbeck-Nordhaus sobre el "ciclo político de los negocios"[10] presumiblemente podría extenderse para cubrir estos casos haciendo endógenas las metas políticas.) Los países recientemente liberados, en donde el desarrollo político está más avanzado que el estado de la economía, enfrentan problemas similares; también aquí puede ser una

necesidad política proclamar tasas de crecimiento o prometer niveles de consumo que son claramente imposibles en esa etapa. Una variante de este dilema es el caso en el cual los planificadores encuentran útil, y quizá necesario, proponer objetivos inaccesibles al plan para asegurar que la economía opere plenamente e incentivar el rendimiento.[11] Puede ser políticamente necesario (o por lo menos posible) establecer como metas estados que son *lógicamente* contradictorios, como cuando los planificadores chinos instruyen a los dirigentes locales para que den prioridad a la industria pero no a expensas de la agricultura, con la esperanza de que este mandato contradictorio establezca una tensión creativa y movilice a las masas.[12] (En realidad, las instrucciones contradictorias parecen haber tenido, con frecuencia, el efecto destructivo del doble vínculo; veremos más al respecto en el capítulo 4.) La maquinaria de la lógica modal puede, sin duda, utilizarse para explorar la lógica de las propuestas políticamente aceptables, pero en lo que sigue presumo siempre que las propuestas son también económicamente posibles. Si ninguna propuesta satisface ambos requerimientos, entonces, el conjunto de estados políticamente posibles es vacío.

Habiendo ofrecido estas observaciones acerca de la interpretación general del modelo, voy a lo específico. Permítasenos comenzar con las modalidades políticas aplicadas al cálculo proposicional. "Existe PP tal que *p*" significa que hay un estado políticamente posible donde rige *p*, y "es políticamente necesario que *p*" significa que *p* rige en *todos* los estados que son PP c.r. al estado dado. Sólo introduciendo el operador de necesidad obtendremos alguna utilidad de este método, ya que la posibilidad en sí misma no nos lleva muy lejos. Como observamos en el capítulo 2, la necesidad habitualmente entra en un modelo a manera de restricción sobre el conjunto de estados posibles, pero con la utilización del método modal hemos visto que puede haber características necesarias no explícitas. Típicamente, hay tres pasos en la construcción de una interpretación modal de los hechos sociales: la especificación de las demandas, la definición de los estados posibles que satisfagan simultáneamente todas las demandas y la identificación de las características comunes a todos los estados posibles. Las características políticamente necesarias pueden adquirir la forma de restricción; es, por ejemplo, virtualmente axiomático que en las modernas sociedades capitalistas la redistribución del ingreso no puede nunca llevarse a cabo por medio de una reducción explícita de los impuestos previos a los salarios nominales de cualquier grupo de ingresos; una distribución más pareja sólo puede lograrse por medio del crecimiento económico, la inflación o los impuestos. Por otra parte, también las necesidades políticas pueden aparecer como el tercer paso en la construcción arriba señalada, como resultado indirecto de las restricciones. En el capítulo 2 di dos ejemplos de dichos rasgos necesarios indirectos: uno, las implicaciones en la tecnología de las demandas simultáneas sobre el consumo y la inversión, y otro, las implicaciones en el plan educacional de las demandas de los padres y los maestros que presionan a las autoridades políticas. Es claro, ahora, cómo pueden surgir nuevos conceptos cuando se confronta el

trabajo de los economistas y los lógicos. Para los economistas, las restricciones son los elementos básicos y el conjunto de estados posibles, una noción derivada; para los lógicos, el conjunto de mundos alternativos a un mundo dado se toma como dado y las características necesarias se derivan de este conjunto. Combinando las dos construcciones de dos pasos con un modelo de tres pasos, surge inmediatamente la diferenciación entre las dos clases de necesidad.

Podemos ir de las nociones primitivas de posibilidad y necesidad a las modalidades iterativas y mixtas. "Es PP que es PP que *p*" significa que hay estados de cosas *x* e *y* tales que *x* es PP c.r.al estado dado, *y* es PP c.r.a *x*, y *p* rige en *y*. Como discutiremos más adelante, esto no necesariamente implica que *p* es PP c.r.al estado dado. La necesidad iterada se interpreta de la manera evidente. Un problema típico en cuestiones políticas es el siguiente. Dado que es políticamente necesario que $\bar{p}$, ¿es PP que es PP que *p* o será más bien el caso que es políticamente necesario que $\bar{p}$? En otras palabras: si una propuesta dada es políticamente imposible ¿es posible alcanzarla haciendo una nueva propuesta? o ¿son tales operaciones de dos pasos todas políticamente imposibles? En cuanto a las modalidades mixtas "es PP que es políticamente necesario que *p*" significa que hay un estado *x* que es PP c.r.al estado dado tal que *p* rige en todos los estados que son PP c.r.a *x*. Incluso cuando es políticamente imposible obtener *p*, puede ser políticamente posible alcanzar un estado desde el cual todos los caminos lleven a *p*. Crear precedentes, realizar inversiones irreversibles que comprometan a las generaciones futuras, explotar complementariedades ocultas o consecuencias imprevistas pueden ser mecanismos pertinentes para este fin, como lo saben bien quienes participan en el proceso presupuestario.[13] De manera similar, "es políticamente necesario que sea PP que *p*" significa que no hay ninguna manera de poner a *p* por fuera del sistema; cualquier cosa que se haga, *p* permanece posible. Los regímenes de derecha parecen incapaces de desterrar el potencial revolucionario, los socialdemócratas, de eliminar la posibilidad de *poujadismo*.

Las modalidades iterativas y mixtas plantean agudamente el problema de cómo se situarán los estados en el *tiempo*. Por razones evidentes, un enfoque de tiempo continuo no puede ser útil aquí, de modo que tendremos que pensar en términos de estados discretos. El caso más simple es el de las elecciones periódicas, pero un análisis más general tiene que tener en cuenta que la mayor parte de la política no presenta este ritmo regular. (Un fascinante problema es pensar las implicaciones para la posibilidad política si se ejecutara la idea —debida a Assar Lindbeck[14]— de tener elecciones distribuidas al azar.) Un "análisis de la situación" que concluya afirmando que una revolución es posible no establece, ni puede hacerlo, un límite de tiempo definido. Por otra parte, la revolución no permanecerá como posible por un período indefinido, incluso suponiendo que no ocurra ningún cambio de estado (definido por un conjunto finito de variables que se mantiene constante a través del análisis). *En política, que nada ocurra*

puede ser un acontecimiento. Este tipo de consideración sugiere una extensión de la noción de estado. Define una *situación* como el estado real más los estados posibles que van con ésta; luego, puede darse un cambio en la *situación* incluso cuando el estado permanece igual. Algunas posibilidades, por así decirlo, tienen una vida corta. De nuevo, esto implica que el tiempo político está puntualizado no sólo por los cambios de estado sino también, generalmente, por el cambio en la situación. En el capítulo 2 observamos que se encontró útil en la investigación de la movilidad social una noción de *inercia acumulativa*; una noción similar —y quizá, todavía más, la idea de una pérdida acumulativa de inercia— puede ser útil en el estudio de los procesos políticos. De nuevo, enfatizo el punto metodológico de que estas nociones no son esenciales aunque pueden ser de utilidad. Redefiniendo el estado por ejemplo, por la adición de nuevas variables, la noción de situación puede siempre convertirse en superflua, de manera que sólo permanezca el estado (como una variable de estado).

Una lógica plenamente desarrollada de la política debe ir claramente más allá de las variables oracionales e introducir la cuantificación. Como ejemplo de enunciado que se presta fácilmente a ser analizado en términos de una lógica modal política con cuantificación podemos tomar el "teorema de la imposibilidad", discutido en el capítulo 1. Puede engañarse a algunas personas todo el tiempo y a todas las personas parte del tiempo, pero no a todas las personas todo el tiempo. En verdad, el problema opuesto es más importante en las sociedades modernas, por ejemplo, la imposibilidad de hacer algo a algunos individuos (ejemplo, el reclutamiento) sin hacerlo a todos. La solución técnicamente óptima de imponer un prolongado servicio militar a algunos individuos, más bien que un servicio corto a todos, puede ser políticamente imposible por las tendencias universalistas inherentes a la sociedad moderna. (Una solución frecuente a este problema es escoger al azar, lo cual reintroduce la universalidad en el nivel de las posibilidades.) En algunos casos puede ser técnica o económicamente imposible otorgar un beneficio a todos al mismo tiempo, mientras que es políticamente imposible otorgarlo a algunas unidades solamente, en cuyo caso la universalidad puede ser un obstáculo para el progreso (a no ser que sea neutralizado por el "efecto túnel" de Hirschman).[15] Un problema similar surge de la división de beneficios cooperativos, cuando todo arreglo factible puede implicar una división desigual. Si todos los grupos sociales se benefician de la estabilidad, mientras cada conjunto de posibles normas estables traería beneficio adicional a algún grupo en particular, es muy posible que no surja ningún arreglo estable. (Téngase en cuenta que esto no es igual al problema del que se corta solo (*free-rider*) en el Dilema del Prisionero. La relación entre estos dos obstáculos para la cooperación se discute brevemente en el capítulo 5.)

Para una visión más detallada de la lógica política cuantificada podemos considerar las siguientes proposiciones:

(1) En una monarquía es necesario que ($\exists x$) (x es un monarca).
(2) En una monarquía es necesario ($\exists x$) (x es necesariamente un monarca).
(3) En una república es necesario que ($\exists x$) (x es presidente).
(4) En una república ($\exists x$) (x es necesariamente un presidente).
(5) Es políticamente necesario que ($\exists x$) (x es un monarca).
(6) ($\exists x$) (es políticamente necesario que x sea un monarca).
(7) Es políticamente necesario que ($\exists x$) (x es presidente).
(8) ($\exists x$) (es políticamente necesario que x sea presidente).

Vimos en el capítulo 1 que la deducción de (2) desde (1) es una inferencia inválida (equivalente de hecho, a la falacia de composición), y de manera similar para los otros tres pares de proposiciones. Así, en 1936 (5) era verdad en Inglaterra, pero (6) era falso. El valor de verdad de las proposiciones (5)-(8) dependerá de las circunstancias específicas a las cuales se apliquen y no diré nada más acerca de ellas aquí, salvo para afirmar que es importante diferenciarlas de las proposiciones (1)-(4), que dependen de una noción totalmente distinta de necesidad. Si estamos considerando la posibilidad y la necesidad lógicas, entonces (1) y (3) son verdaderas, mientras que (2) y (4) son falsas. Por otra parte, parece haber un sentido, si sólo pudiéramos acceder al mismo, en que (2) difiere esencialmente de (4). Los reyes son reyes por algún tipo de necesidad y no todos pueden llegar a ser reyes; los presidentes son presidentes más o menos por accidente, y casi todos pueden llegar a ser presidentes. Permítaseme proponer una exploración tentativa (y levemente frívola) de estos temas.

Una manera de mostrar la diferencia entre (2) y (4) fue propuesta recientemente por Saul Kripke y refinada aun más por Michael Dummett.[16] Defendiendo un enfoque *genético* de la necesidad, Kripke argumentó que las propiedades *metafísicamente necesarias de un objeto* estén relacionadas con el camino por el cual llegó a ser o con el material del cual está hecho. Uno de sus ejemplos, conectado a (2), es, por ejemplo, su discusión de si sería verdadero decir que la reina Isabel podría no haber nacido de sangre real. Contesta afirmativamente a esta pregunta ya que una respuesta negativa tendría que demostrar la necesidad de que los padres de la reina fueran de familia real, y esto, de acuerdo con Kripke, es un hecho contingente, presumiblemente por la posibilidad de cambio de bebés, etcétera. Lo que *sí* contesta negativamente, sin embargo, es la cuestión de si la reina Isabel podría haber nacido de padres distintos a los reales. El hecho de tener como padres a sus padres reales es una propiedad necesaria (según el criterio genético) de la reina Isabel, mientras que el hecho de tener a John F. Kennedy por hermano no sería (bajo el mismo criterio) una característica necesaria de Edward Kennedy.

Utilizando este criterio genético de necesidad, las posibilidades de cambios de bebés o de reyes elegidos hacen que la proposición (2) sea falsa. Sin embargo, puede argumentarse que estos son casos perversos o

marginales, que no podrían tornarse típicos sin que la noción misma de reinado se disolviera. Por "típico" quiero significar aquí "como un principio" y no sólo "frecuente". Nada excluye la posibilidad de una conspiración intertemporal de comadronas para hacer cambios de bebés de familias reales, pero esto no sería parte del concepto de rey. Si las cosas se dieran siempre tal como ocurren típicamente, tendríamos que concluir que (2) se haría verdadero. Por otra parte, también (4) podría no ser verdadero bajo una cláusula semejante; no es diferente el valor de verdad de esta proposición si descartamos la posibilidad de un impostor que tome el lugar del presidente electo, etcétera.

Antes de aceptar la verdad de la proposición (2) en el caso corriente, deberíamos mirar más de cerca las restricciones impuestas a los individuos y el valor de verdad de las proposiciones que adscriben propiedades a objetos no existentes. Si todos los mundos posibles tienen como miembros exactamente a los mismos individuos, entonces la reina Isabel es reina en todos ellos y, por lo tanto, necesariamente reina en el nuestro. Sin embargo, si existe un mundo donde la reina Isabel tenga un hermano o hermana mayor, entonces no es necesariamente reina en ese mundo y, por lo tanto, no es necesariamente reina en el nuestro. Finalmente, si existe un mundo donde la reina Isabel no esté, pero ningún mundo donde tenga un hermano o hermana mayor, entonces debemos ir a la interpretación de las proposiciones sobre objetos no existentes, de las que ya dijimos algo en el capítulo 1. Si todas las proposiciones acerca de individuos no existentes se toman como verdaderas, entonces es verdad que la reina Isabel es reina en los mundos en los que ella no existe y es necesariamente reina en nuestro universo. Si, de otro lado, adscribimos falsedad a las afirmaciones sobre no entidades, entonces la reina Isabel no es reina en los mundos donde ella no existe y, por lo tanto, no es reina en nuestro mundo. Ser rey o reina puede parecer una propiedad esencial por su articulación con la propiedad esencial de tener los padres de uno por padres, pero en una mirada más detenida se torna accidental, porque no es esencial para los padres tener a sus hijos como hijos. El criterio genético de necesidad funciona solamente hacia atrás en el tiempo. Incluso cuando abstraemos el cambio de bebés en el mundo real, el cambio de bebés entre los mundos destruye la necesidad de ser reina. Lo que se preserva es, sin embargo, la necesidad de la realeza, es decir, la propiedad de pertenecer a una familia real, por lo menos si suponemos que la atribución de esta propiedad a la reina Isabel es verdadera también en los mundos en los cuales ella no existe.

De estas graciosas observaciones paso ahora a cuestiones más importantes: las propiedades lógicas que podemos atribuir a la relación PP entre mundos. En el capítulo 1 mencionamos que diversos sistemas de la lógica modal (tales como S-4 o S-5) pueden ser clasificados de acuerdo con las propiedades que le adscriben a esa relación. Argumentaré que los sistemas de la filosofía política o, por lo menos, parte de ellos pueden ser formulados en términos de dichas propiedades. Hay un número indefinido

de propiedades que pueden ser elegidas a fin de inspeccionarlas, pero me limitaré a siete propiedades discutidas en un texto reciente.[17]

(l) *Existencia*: $(\forall x)(\exists y)$ (y es **PP** c.r.a x)

Ya me he referido a algunas razones por las que ésta no es una propiedad trivialmente satisfecha. El conjunto de estados políticamente posibles puede ser vacío, porque todos los posibles estados económicos son políticamente inaceptables. Incluso si un estado dado es políticamente aceptable y se torna realizable, esto no prueba que fuera políticamente posible *ex ante*, porque puede haber sido realizado por accidente más bien que por una acción deliberada e inteligente. (Consúltense los comentarios acerca de la definición apropiada de poder en la primera parte de este capítulo.) En una sociedad donde los cambios en las personas y en las estructuras políticas dependen de algún mecanismo de azar con alguna probabilidad desconocida de distribución, como en el cuento de Borges "La Lotería", nada podría ser realizado por los actores mismos, aunque cada estado sería seguido por algún otro estado. Menos imaginativa, la no existencia puede resultar ser una contrafinalidad (en el sentido que será definido en el capítulo 5). Si los actores $a_1,a_2...a_n$ se proponen producir los estados $s_1,s_2...s_n$ respectivamete, y si por cada i es el caso que s_i sólo puede realizarse bajo la condición de que por lo menos una $j \neq i$, s_j no sea propuesta, entonces ninguno de los estados se realizará. Esta situación podría surgir si hay diversos grupos minoritarios con intereses diferentes en cada uno de los (abstractamente posibles) estados. China, al final de la dinastía Han, aporta un ejemplo de esa estructura caótica, donde la carencia de un contexto político común, la intensidad del conflicto y la multiplicidad de intereses hicieron la acción política exitosa —en el sentido de alcanzar un resultado definido con anterioridad— virtualmente imposible.[18] Algunos de los actores alcanzarán lo que deseaban, en el sentido de obtener el poder, pero la utilización a la que el poder quede atado puede ser radicalmente diferente de la que ellos habían pensado.

(2) *Reflexividad limitada*: $(\forall x)((\exists y)$ (x es **PP** c.r.a $y) \supset (x$ es **PP** c.r.a $x))$

Si un estado dado es en sí mismo PP c.r.a algún estado, entonces es también PP c.r.a sí mismo. O: si un estado puede ser alcanzado, entonces, también puede ser mantenido. O de nuevo: si la conservación del *status quo* no es políticamente posible, entonces ese estado no puede ser *alcanzado* desde ningún estado (incluso si, como ya fue explicado, *sobreviene* de manera no intencional y no inteligente). Ciertamente, esta propiedad no rige universalmente. En las sociedades modernas industrializadas, la inmovilidad política es imposible y, sin embargo, ningún estado dado (cuya permanencia es imposible) resulta de la realización deliberada de posibilidades *ex ante*. Esta moderna preocupación por el cambio fue expresada, de manera admirable, por Leibniz: "Regresamos si no progresamos porque es imposible detenerse."[19] Podemos aquí hacer una pausa para observar que

en la noción de reflexividad limitada hay una idea implícita de *estados inherentemente inaccesibles*, los cuales conforman el concepto polar, por así decir, de los estados cuyo conjunto de estados políticamente posibles es vacío. Independientemente de la cuestión de si se da la reflexividad limitada, podemos preguntarnos si hay estados políticos que puedan realizarse sólo de manera no intencional o no inteligente, así como puede haber estados que sólo puedan permanecer de esa manera. Podemos preguntarnos, por ejemplo, si podría haber *propuestas autodestructivas*, es decir metas cuya enunciación es incompatible con su realización. La propuesta de lograr un estado de inocencia podría ser candidata a ese tipo de metas; compárese también la discusión de la simetría que daremos posteriormente.

(3) *Reflexividad*: $(\forall x)$ (x es **PP** c.r.a x)

A fortiori, esto no es así cuando una de las dos propiedades precedentes no se da.

(4) *Transitividad*:$(\forall x)$ $(\forall y)$ $(\forall z)$ ((y es **PP** c.r.a x y z es **PP** c.r.a y) $\supset$ (z es **PP** c.r.a x))

Que PP c.r.a es un estado que es PP c.r.al estado real, es en sí mismo PP c.r.al estado real. Esta es una proposición muy controversial, y puede decirse que su aceptación es la esencia de las aproximaciones utópicas a la política. El utopismo dice que lo que puede hacerse en dos pasos también puede ser hecho en un paso. Me limito aquí a los aspectos políticos del utopismo. La mayor parte de las personas aceptarán que en las cuestiones concernientes al capital y al crecimiento económico frecuentemente es imposible tomar atajos, pero es menos evidente que restricciones similares puedan regir para el desarrollo político.

La refutación del método utópico puede darse sobre dos bases diferentes. Primero, uno puede argumentar sobre la analogía del horizonte espacial: para ver lo que hay atrás de la próxima colina primero hay que llegar allí. De hecho, la intransitividad y la asimetría del ver han sido explotadas en la construcción de los modelos semánticos de la lógica modal.[20] Más profundamente, Ernst Bloch ha utilizado la idea de un horizonte para expresar el concepto de una *utopía relativa*.[21] Más profundamente aún, podemos apelar a la fenomenología de Husserl, con sus diferenciaciones entre las intenciones llenas y vacías. Desde mi punto de vista actual tengo una percatación total del frente de todos los objetos en mi campo visual, pero sólo una percatación formal y vacía del lado oculto de mi visión. Sé que los objetos tienen que tener un aspecto posterior, pero ignoro cómo esa parte puede aparecer detalladamente. También, desde mi punto de vista, sé que si me desplazara a otro lugar en este campo, nuevos objetos serían visibles para mí, incluso si no los conociera. Lo mismo es válido para mi conciencia del futuro, que se da en "protenciones" vacías y formales, las que junto con mis impresiones actuales y mis retenciones del pasado constituyen la red del tiempo interno subjetivo. La estructura formal

de esta red impone en toda experiencia presente el sello indeleble de que es también algo que "habrá sido".[22]

De manera similar, puedo ignorar qué estados son PP c.r.a los estados que son PP c.r. al presente, pero sé —de una manera vacía y general— que cualquier estado que llegue a realizarse tendrá también su horizonte de posibilidades políticas. Y si uno no puede saber cuáles estados son posibles en dos movimientos, entonces, *a fortiori*, uno tampoco puede luchar para realizarlos. Así, la transitividad de la relación PP se rompe. La noción de la ignorancia de las posibilidades futuras puede ser tomada aquí en un sentido débil o fuerte. O bien puede considerarse que significa que no conozco cuál de los estados, en un espectro dado y constante de estados abstractamente posibles, aparecerá como PP c.r.a los estados que son PP c.r. al presente o que implica que, de acuerdo con lo que sé, algunos estados cualitativamente nuevos pueden adicionarse al espectro. En mi opinión, la útima interpretación es la más interesante. Me parece que también está en la raíz de la crítica de Marx al utopismo, en un pasaje de *La guerra civil en Francia*:

> La clase obrera no esperaba milagros de la Comuna. No tiene utopías prefabricadas para introducir *par décret du peuple*. Sabe que para lograr su propia emancipación, juntamente con la forma más alta a la cual la sociedad actual irresistiblemente tiende por medio de sus propios agentes económicos, tendrá que atravesar una serie de procesos históricos que transformarán las circunstancias y a los hombres. No tiene ningún ideal para realizar, excepto liberar los elementos de la nueva sociedad de los que está cargada la vieja sociedad burguesa decadente.[23]

Segundo, se puede argumentar a un nivel más táctico que, incluso si los dirigentes políticos tuvieran el conocimiento pleno de lo que serán PP c.r.a los estados que son PP.c.r. al presente , sus seguidores podrían carecer de la madurez política para saltar el paso intermedio. Este argumento es realmente más plausible si se utiliza en combinación con el primero, sobre la base de una distinción entre los pioneros y los que llegan tarde al desarrollo político. Para el país que es el primero en desarrollar, digamos, las instituciones democráticas o las organizaciones de la clase trabajadora, cada nuevo avance se da en un territorio desconocido y agrega nuevas posibilidades al espectro. Aquí, el argumento de los párrafos precedentes es suficiente para refutar al utopismo. Para los que llegan más tarde, sin embargo, el ejemplo pionero demuestra los estados que serán posibles en uno o dos movimientos. Para esos países habrá una tentación permanente de obviar el paso intermedio, para ir directamente, digamos, del feudalismo al comunismo o para importar instituciones democráticas ya totalmente desarrolladas. Contra esto, algunos argumentarán que el conocimiento abstracto no es suficiente para crear las condiciones de formas más avanzadas, lo cual puede ser viable sólo si está enraizado en la experiencia concreta, que tiene que ser adquirida gradualmente. A la frase de Marx "El

país más desarrollado industrialmente sólo muestra a los menos desarrollados la imagen de su propio futuro"[24] podemos oponer el ácido argumento de Trotsky: "Para tales pedantes la historia de una nación capitalista repite la historia de otra, desde luego, con algunas diferencias, más o menos importantes."[25]

Para los dirigentes políticos que han adquirido alguna comprensión de qué será posible en dos o más movimientos, surge una paradoja peculiar, encarnada en la *teoría del fracaso necesario.*[26] Según ésta, los dirigentes políticos podrían proponer metas que, saben, son políticamente imposibles, porque sólo a través de la dura derrota en la lucha por esas metas, la clase revolucionaria adquirirá la madurez y organización requeridas para su realización final. Esto puede ser llamado una profecía indirectamente autocumplida, en contraste con las profecías inmediatamente autocumplidas, discutidas con anterioridad en este capítulo. Debería enfatizarse que, desde un punto de vista moral, la teoría del fracaso necesario es mucho más aceptable si se la ve como una explicación posterior de los sucesos y no si se la hace fundamento de un análisis anterior al hecho. *En 1850*, Marx dijo que

> Con la excepción de unos cuantos capítulos cortos, cada parte importante de los anales de la revolución de 1848 a1849 lleva el encabezamiento: ¡Derrota de la revolución! Pero lo que sucumbió en esas derrotas no fue la revolución. Fueron los accesorios prerrevolucionarios tradicionales —personas, ilusiones, conceptos, proyectos— surgidos de las relaciones sociales que aún no habían llegado al punto culminante de los antagonismos de clase y de los cuales el partido revolucionario antes de la revolución de febrero no estaba libre ni de los cuales podía liberarse por la victoria de febrero, sino sólo por una serie de derrotas.[27]

Por lo que yo sé, en sus escritos *anteriores* a 1848, sin embargo, Marx nunca enfatizó esta idea. La idea manipuladora de que la clase obrera puede avanzar sólo luchando por metas que los dirigentes han elegido deliberadamente porque están fuera de alcance es propia de Rosa Luxemburgo.[28] Es claramente imposible decirles a los trabajadores que van a alzarse para ser derrotados y que, por medio del desarrollo, alcanzarán la conciencia para ser una clase revolucionaria madura, porque los trabajadores dispuestos a aceptar ese argumento ya poseerían esa conciencia. Las alternativas para los dirigentes son, *o bien* renunciar a la dirección, para que los miembros menos experimentados puedan cometer *sus propios* errores, *o bien* presentar a las bases la victoria como una posibilidad real, aunque sepan que la derrota es segura. La primera opción, prácticamente no necesita ningún comentario; la segunda es manipulación, aunque no de la variedad leninista. En lo que sé, Lenin nunca enfatizó la dialéctica en la cual el fracaso se transformaba en éxito, y la falsedad, en verdad. El fracaso puede haber sido para él, como lo fue para Rosa Luxemburgo, el resultado de una falta de conciencia de clase, pero no —como lo fue para ella— el prerrequi-

sito de la conciencia de clase. (Es muy fácil ver que muchos de estos comentarios son también válidos para la relación entre maestro y alumno, padre e hijo o psicoterapeuta y paciente; compárese, por ejemplo, la utilización de instrucciones paradójicas en la terapia del comportamiento. Por otro lado, hay diferencias importantes entre el caso colectivo e individual. Este último tiene, notablemente, mucha más posibilidad de autoeducación estratégica.)

(5) *Simetría limitada*: $(\forall x)\ (\forall y)\ ((x \text{ es } \mathbf{PP} \text{ c.r.a } x) \supset ((y \text{ es } \mathbf{PP} \text{ c.r.a } x) \supset (x \text{ es } \mathbf{PP} \text{ c.r.a } y))$

Si un estado dado es PP c.r.a sí mismo, y si además una alternativa *y* es PP c.r.a ese estado, el estado es también PP c.r.a *y*. O: si el *status quo* es PP, entonces todo cambio es reversible. Esta es una proposición que no parece evidentemente verdadera ni evidentemente falsa, así que tendríamos que mirar la evidencia empírica para circunscribir el tipo de situaciones donde probablemente se dé.

(6) *Simetría*: $(\forall x)\ (\forall y)\ ((y \text{ es } \mathbf{PP} \text{ c.r.a } x) \supset (x \text{ es } \mathbf{PP} \text{ c.r.a } y))$

Cada estado es PP c.r.a los estados que son PP c.r.a sí mismos. O: todo cambio político es reversible. Esto es indudablemente incorrecto. Por ejemplo, la simetría verbal de los conceptos de revolución y contrarrevolución es sólo aparente, porque la meta de los contrarrevolucionarios no es verdaderamente lograr una situación en la que la revolución se mantenga a la orden del día. La actual (diciembre de 1976) situación en Chile es un retroceso *más allá* y no *a* la situación previa a Allende, como muy tarde lo descubrió Eduardo Frei. Para una expresión explícita de esta idea no podemos hacer nada mejor que citar una declaración de Valery Giscard d'Estaing: *Il n'est certainement pas question de revenir à la situation d'avant 1968, et d'abord parce que la situation avant 1968 comportait les conditions qui ont créé 1968.*[29] La irreversibilidad política puede también surgir de la imposibilidad de cortar sus pérdidas, aunque incluso sea racional hacerlo. El proyecto Concord y la guerra de Vietnam son apenas dos ejemplos de empresas que nunca deberían haberse iniciado y se hicieron virtualmente irreversibles por razones de prestigio. La experiencia, una vez adquirida, puede ser difícil de deshacer. Así, puede conjeturarse que los países tradicionalmente neutrales como Suiza o Suecia no retornarían a la neutralidad si un día se vieran envueltos en una guerra. En relación con esta fuente de asimetría, anotemos que el problema de la irreversibilidad puede volcarse a la forma de definición de los estados. Por lo tanto, si B es PP c.r.a A y si tener un pasado sin B es parte de la descripción de A, entonces A no puede ser PP c.r.a B. Si, por otro lado, los estados son descritos solamente en términos de variables de estado, entonces la irreversibilidad no puede ser afirmada sobre esas bases puramente lógicas. Si la inocencia se define como un estado que nunca ha conocido la duda, entonces la inocencia, una vez perdida, no puede nunca ser recuperada. Si, por otro lado, la inocencia

es definida como una cierta forma de comportamiento, la cuestión de la irreversibilidad es empírica más bien que puramente conceptual.

Tomaré estas últimas observaciones como punto de partida de algunos comentarios sobre la noción hegeliana y marxista de *negación de la negación*. No suscribo ninguna "ley" de la negación de la negación como fue postulada por Engels, si por ello se implica que *todos* los procesos de cambio en la naturaleza, la conciencia o la sociedad proceden de acuerdo con este criterio. Por otro lado, no estoy totalmente de acuerdo con los críticos del marxismo que dicen que "la principal dificultad de esta ley de la negación es que puede hacerse coincidir con casi cualquier cosa escogiendo cuidadosamente lo que puede tenerse en cuenta como término de negación".[30] Es indudable que muchas instancias de la "ley" son simples juegos verbales, en los que la mano no es suficientemente ligera para engañar al ojo más crédulo. Sin embargo, pienso que muchos de los ejemplos tradicionalmente subsumidos bajo el encabezamiento de "negación de la negación" muestran un patrón que puede definirse con alguna precisión y ser aplicado en muchos contextos.

Un ejemplo sencillo del tipo de criterio que tengo en mente es aportado por Husserl en su discusión sobre la *duda* como una modalidad de la conciencia.[31] Aquí, él anota que la duda puede ser vista como la negación de la creencia ingenua y que, incluso después de restablecerse la creencia, la duda deja su marca irreversible en la conciencia. El paso de la creencia no pensada a la duda puede ser visto como algún tipo de "negación", como también el paso de la duda a la creencia madura; evidentemente, el resultado final de estas dos operaciones no es el mismo estado que el del inicio. El paso desde los *Cantos de la inocencia a los Cantos de la experiencia* puede verse como una negación, y el paso desde estos últimos al estado mental capaz de escribir ambos como una segunda negación. Donne nos dice que "Es menor la pesadumbre de ser malo que la de haber sido honorable" y por lo menos tenemos que estar de acuerdo en que estos dos estados son totalmente diferentes. Emily Dickinson hace una exposición perfectamente explícita de esta idea:

> No puede experimentar la escasez
> Quien la Abundancia no haya conocido,
> El hecho de la Hambruna no podría ser
> Sin el Hecho del Maíz,
>
> La Carencia es un Arte mezquino
> Adquirido por un Revés,
> La Pobreza que no fue Riqueza
> No puede ser Indigencia.[32]

Subyacente en todos estos casos a los que el lector, sin duda, podrá añadir unos cuantos más, señalamos la siguiente estructura. Dada una secuencia real *p-q-r*, decimos que ejemplifica el modelo de la negación de

la negación si (i) *p,q* y *r* son *contrarios*, en el sentido que ningún par de ellos puede ser realizado conjuntamente; (ii) la secuencia *p-q-p* es imposible; (iii) la secuencia *p-r* es imposible. La primera condición da un significado (más bien débil) a la noción de una negación. Tiene el efecto de excluir el crecimiento puramente órganico y acumulativo, en que cada etapa incluye directamente a la anterior, como (digamos) el proceso de aprender un idioma extranjero. La segunda condición impone antisimetría, significando que la experiencia de ir a través de *q* cambia irreversiblemente al sujeto del proceso. La tercera condición impone antitransitividad, sugiriendo que el paso a través de *q* es indispensable para llegar a *r*. Sostengo que los procesos que satisfacen (i)-(iii) forman una subclase importante de estructuras en general cambiantes; también pienso que corresponden bastante bien a los ejemplos dados por Hegel y Marx. En particular, la secuencia feudalismo-capitalismo- comunismo[33] muestra todas las características anteriormente especificadas. En este importante ejemplo, la imposibilidad en cuestión es en parte económica, en cuanto que se requiere el capitalismo para el desarrollo de las fuerzas productivas, y en parte es política, en cuanto que la alienación de la sociedad capitalista es un prerrequisito para el desarrollo total de la libertad bajo el comunismo. Los hechos de los últimos años en Portugal y Chile pueden servir como ejemplos de la negación de la negación, en contradiferencia con Grecia, donde el paso por el régimen de los coroneles no parece haber significado mayor diferencia. Tengo que admitir que estos conceptos no son rigurosos, pero los encuentro sugerentes.

(7) *Cuasi-equivalencia*: $(\forall x)\ (\forall y)\ (\forall z)\ ((y \text{ y } z \text{ son } \mathbf{PP} \text{ c.r.a } x) \supset (y \text{ es } \mathbf{PP} \text{ c.r.a } z \text{ y } z \text{ es } \mathbf{PP} \text{ c.r.a } y))$

Todos los estados que son **PP** c.r.a un estado dado son **PP** c.r.a cada uno. Esto puede ser incluido en la definición de un régimen parlamentario, porque, en ausencia de la cuasi-equivalencia, el partido perdedor en una elección podría tener que retirarse, que es lo que por supuesto tiene que hacer con frecuencia el lado perdedor en una revolución. Tenemos que anotar que si el conservatismo es **PP**, entonces la cuasi-equivalencia implica simetría.

No propondré más propiedades de la relación **PP**,[34] pero concluiré el capítulo con algunas otras observaciones sobre la argumentación general que he presentado. En primer lugar, quisiera volver a destacar los modestos objetivos del enfoque. Pienso que es necesario por el aspecto obviamente novedoso de aplicar la lógica modal a los asuntos políticos; si no quiero terminar en el Rincón de lo Seudo de alguien, debo repetir que pretendo ser útil más bien que profundo. Los mundos posibles son sólo una manera de hablar y, en principio, nada puede decirse con los conceptos modales que no pueda también ser dicho sin ellos. Sin embargo, para algunos propósitos el lenguaje de los mundos posibles puede ser más conveniente y menos fastidioso que formas alternativas de expresión. He

utilizado varias veces una analogía que destaca bien este punto, por ejemplo la comparación con la histéresis. En general, sabemos que el pasado no puede tener impacto causal sobre el presente (por encima y a través de la influencia mediada de las huellas dejadas por el pasado en el presente), sin embargo podemos encontrar conveniente hablar como si el pasado ejerciera algún tipo de influencia directa en el presente. De manera similar, también sabemos *a priori* que una posibilidad no es el tipo de cosa que puede ejercitar influencia causal o que dos estados no pueden diferenciarse *sólo* por el halo de sus posibilidades. No obstante estas verdades elementales, podemos encontrar que es útil describir una situación como un estado con sus posibilidades adjuntas, más bien que como algo totalmente real. Los valores pasados o posibles de las variables, pueden servir algunas veces como sustitutos de los valores reales y presentes.

En segundo lugar, el lector puede haberse preguntado cuáles *son* exactamente los usos de las modalidades políticas. De lo dicho anteriormente surgen dos aplicaciones fundamentales. La primera es que podemos clasificar los regímenes políticos de acuerdo con las propiedades que imponen a la relación de posibilidad política; esto tiene vigencia tanto para los regímenes reales como para los regímenes posibles, que pueden generarse por una combinación de diversas propiedades, tales como las siete ya mencionadas. El lector puede observar que esto implica que las transiciones políticamente posibles pueden darse sólo en el interior de los regímenes y nunca entre un régimen y otro. Al escogerse este método, tratamos con los aspectos más rutinarios de la política, aunque esto no implica que nos restrinjamos a la noción que tiene el actor de lo que es y no es posible. La segunda aplicación es en el análisis de las teorías del desarrollo político, en el que la lógica y la consistencia de un esquema conceptual pueden ser puestas a prueba en la forma que se bosquejó en el capítulo. El análisis de la negación de la negación aporta un ejemplo de cómo puede aclarar nociones notablemente evasivas. Esta interrelación entre conceptos políticos bien puede ser presentada de esta manera: si, por ejemplo, un autor afirma la existencia, transitividad y simetría, pero niega la reflexividad de la relación **PP**, entonces está en problemas.

Notas

1. Elster (1976a); también Davidson (1973), citado en el capítulo 1.

2. Véase Simon (1953) para una buena exposición del enfoque "del poder como causalidad".

3. Russell (1938).

4. Goldman (1972, pág.223).

5. Véase Lave y March (1975, págs. 294 a 305).

6. Chomsky (1964, pág.59). Para una diferenciación similar (y explícitamente spinozista) entre *la parole parlée y la parole parlante* véase Merleau-Ponty (1945, pág. 229).

7. Sartre (1943, pág.58).

8. Puede ser válido explicar por qué esto es así. Supongamos que un dirigente político o un científico social piensa que el resultado de una situación es una función de variables x_1, x_2,...x_n y de su misma predicción (y) de su resultado: $z=f(y, x_1, x_2 \ldots x_n)$. En ese caso, él puede buscar o bien un punto fijo $\bar{z}$ de manera que $\bar{z}=f(\bar{z}, x_1, x_2 \ldots x_n)$ o puede buscar el valor de $\underline{z}$ que maximiza alguna función $g(z)$ del resultado, pero normalmente no puede hacer ambas sin que el sistema se vuelva sobredeterminado (Haavelmo, 1971, pág.46). El National Economic Bureau no puede pronosticar los precios correctamente teniendo en cuenta a la vez el impacto de la predicción sobre los agentes económicos y elegir la predicción para maximizar el nivel de empleo en el año siguiente. El dirigente revolucionario puede buscar predecir, por ejemplo, qué es lo que tiene un impacto máximo sobre la combatividad de la clase revolucionaria, pero en este caso tiene buenas posibilidades de equivocarse; o puede tratar que su análisis de la situación se incorpore como uno de los elementos de la situación que analiza, pero en ese caso no hay motivo para pensar que su causa será seguida.

9. Véase también la nota 43 del capítulo 5.

10.Véase Kalecki (1943), Nordhaus (1975) y Lindbeck 1976) para una exploración de este concepto.

11. Heal (1973, pág.20).

12.Véase Schurmann (1967, págs.76 y siguientes) y Eckstein (1968, pág. 713) para la utilización de las directivas contradictorias por los políticos chinos. Para la discusión de las fluctuaciones generadas por este estilo político véase Skinner y Winckler (1969), Suttmeier (1974), Nathan (1976) y Winckler (1976).

13. Wildavsky (1964, pág.111).

14. Lindbeck (1976, pág.18 n.8).

15. Hirschman (1973, pág.545) describe el efecto en estos términos: "El bienestar de un individuo depende de su estado actual de satisfacción (o del ingreso, como un sustituto) al igual que de las expectativas de satisfacción (o ingreso). Supóngase que el individuo tiene poca información acerca de su ingreso futuro, pero en algún momento algunos de su familiares, vecinos o conocidos mejoran su posición económica o social. Ahora, él tiene algo por qué guiarse: esperando que llegue su turno en su debido momento, obtendrá gratificación del avance de los demás... por algún tiempo."

16.Kripke (1972) y Dummett (1973, págs.130-131).

17. Snyder (1971, pág. 210). Las propiedades son identificadas por Snyder debido a su importancia para los sistemas que estudia, y no hay razón por la que deberían ser especialmente importantes para las modalidades políticas. Se las examina aquí simplemente para ilustrar el poder expresivo del lenguaje modal; además, propiedades como la existencia, la reflexividad, la transitividad y la simetría tienen una simplicidad que las hace candidatos obvios para el análisis.

18. Balacz (1968, pags. 71 - 138)

19. *Regredimur nisi progrediamur, quia stari non potest* (Leibniz, 1948, pág. 94). Creo que esta proposición expresa de una manera admirable una diferencia específica que separa el mundo moderno del tradicional. Leibniz probablemente intentó que cubriera muchos niveles, desde el mundo en su conjunto hasta la experiencia de la conciencia individual, pasando a través de la evolución de la sociedad y el estado. Para el lector interesado menciono aquí algunos problemas que surgen si interpretamos la proposición como si fuera acerca del placer individual; véase especialmente Leibniz (1875-90, vol.V, págs 175, 182). Si suponemos que la utilidad es una función continua del nivel de consumo y de la tasa de cambio de consumo, tal que $u(t)=f(c(t), dc/dt)$ entonces se puede interpretar que Leibniz postuló: i) que ninguna utilidad se deriva del consumo constante, es decir que $f(c,o)=0$ para todo c, y ii) que ninguna corriente de consumo puede dar una corriente constante de utilidad, de modo que la ecuación diferencial $f((c(t), dc/dt= k$ no tiene solución para k positiva. Aanund Hylland ha demostrado (en una nota al autor) que si f es una función no decreciente del primer argumento y una función creciente del segundo, entonces la ecuación diferencial siempre tiene solución, y que si es no decreciente en ambos argumentos tiene una solución en el sentido de que existe casi en todas partes una función diferenciable que satisface la ecuación diferencial en casi todas partes. Si permitimos que la utilidad sea una función decreciente del consumo, podemos

encontrar una función f tal que la ecuación diferencial siguiente no tenga solución, pero ¿hubiera Leibniz aceptado esta condición?

20. Hughes y Cresswell (1972, págs. 63 y siguientes).

21. Bloch (1954).

22. Véase Husserl (1966) para el problema de la temporalidad.

23. Marx (1871, págs. 96-97).

24. Marx (1867, págs. 8-9).

25. Trotsky (1972, pág. 50),

26. Tomo esta frase de Avineri (1969, pág. 239), quien parece, sin embargo, usarla en el sentido de un fracaso *inevitable* (dadas las causas prevalecientes), mientras que yo la uso en el sentido de un fracaso *indispensable* (dado el objetivo a ser conseguido).

27. Marx (1820, pág. 33).

28. Luxemburgo (1899, págs. 404-405); Luxemburgo (1906, págs. 133, 141, siguientes).

29. Entrevista en *Le Monde*, 8 de enero 1975.

30. Acton (1967).

31. Husserl (1967, págs. 33 y siguientes).

32. *Complete Poems*, No.771; véase también No.1197 y 1515. El verso de Donne está en su elegía *The Anagram*. La misma distinción, pero la conclusión opuesta, se expresa en "Es mejor haber amado y perdido que no haber amado nunca."

33. Marx (1867, pág.763), Véase también Marx (1862-63, vol. 2, pág.111; vol. 3, pág.41).

34. Además de la reflexividad limitada podríamos examinar la siguiente propiedad: $(\forall x)\,((\exists y)\,(y \text{ es } \mathbf{PP} \text{ c.r.a } x) \supset (x \text{ es } \mathbf{PP} \text{ c.r.a } x))$. Escribiendo "$\mathbf{PP}(x)$" para el conjunto de estados que son **PP**c.r.a un estado x, podríamos preguntar si la relación **PP** está completa en $\mathbf{PP}(x)$, que sería una forma menos estricta para la cuasi-equivalencia. O podríamos preguntar si el conjunto $\mathbf{PP}(x)$ contiene cualquier (y de ser así cuántos) subconjunto tal que la relación de accesibilidad sea una relación de equivalencia cuando está restringida a un subconjunto dado; la existencia y el número de tales conjuntos de estados podría arrojar luz sobre el régimen en cuestión. Muy posiblemente, los resultados de la teoría de grafos podrían también ser invocados aquí. Es una cuestión que necesita mayor reflexión decidir si esta u otras preguntas tienen un interés sustancial.

4

Contradicciones de la mente

Este capítulo y el siguiente tienen un objetivo común que se podría llamar la "reivindicación de la dialéctica". Me apresuro a decir que no deseo reivindicar todas las modalidades de razonamiento que amigos y enemigos han llamado "dialécticas". Al contrario, pienso que los tratados sobre dialéctica escritos por Hegel, Engels y Mao Tse-tung a duras penas ofrecen alguna idea interesante. Lo comprensible es trivial; lo que parece ser profundo es confuso, no sólo en una primera, sino en la décima o centésima lectura. En vez de intentar hacer comprensible lo que estos y otros escritores han dicho *acerca* de la dialéctica, intentaré hacer explícita la estructura lógica (que podría llamarse dialéctica) de los escritos no metodológicos de Hegel y Marx. Me centraré en particular en el concepto de contradicción. Soy tan incapaz como Karl Popper de comprender lo que Hegel dice en su *Lógica* acerca de la contradicción, pero pienso que es posible hacer comprensible la utilización que hace del término "contradicción" en la *Fenomenología del espíritu*, que sin comparación alguna es, para mí, su más grande trabajo. Asimismo, pienso que las más importantes manifestaciones del término[1] en los escritos de Marx se prestan a ser interpretadas en los términos de la lógica cotidiana. Argumentaré finalmente que estos conceptos están no sólo perfectamente claros, sino que también son extremadamente fértiles.

No reconozco distinción alguna entre el entendimiento (*Verstand*) limitado al razonamiento estéril de la lógica formal y el razonamiento (*Vernunft*) que procede con el método creativo de la dialéctica. Tampoco estoy dispuesto a aceptar que mi rechazo de esta diferenciación demuestra simplemente que no es válida. No creo que la relación entre el pensador dialéctico y el analítico pueda compararse a la relación entre lo sano y lo insano, donde el primer miembro de la pareja pueda entender al segundo, pero no viceversa. Tampoco acepto que la superioridad del pensador dialéctico se demuestra en que él es capaz de resolver problemas que ni existen como tales para la mente analítica. Finalmente, creo que si uno está

en lo correcto podría convencer a otros por medio de la *argumentación* más bien que exigiendo una *conversión* espiritual.[2] Por otra parte, es difícil refutar a quien insiste en que todas las objeciones que se hacen son sólo otro síntoma de que no se ha entendido nada. De ahí la necesidad de una estrategia indirecta: como todos los pensadores dialécticos están de acuerdo en que análisis como el de Hegel sobre el amo-esclavo o la teoría de Marx acerca de la tasa decreciente de ganancia son el clímax del razonamiento dialéctico, el ponerlos en un lenguaje analítico común daría *prima facie* una razón para pensar que la brecha no es tan grande.

En este capítulo haré primero una breve exposición de la *teoría* de las contradicciones de Hegel (diferente de su *utilización* del concepto), llamaré la atención sobre sus principales fuentes y comentaré las críticas analíticas comunes a la teoría. Me esforzaré por demostrar que esta crítica, aunque válida, es más bien superficial y que es posible definir una noción de "contradicción" que no esté expuesta a esas objeciones. Esta noción es mayormente explorada en dos variantes principales: las contradicciones de la mente, que nos ocuparán en lo que resta de este capítulo, y las de la sociedad, que son el objeto del capítulo 5.[3] Las estructuras mentales que se dicen contradictorias se subdividen en dos clases principales: creencias contradictorias y deseos contradictorios.[4] Puede ser útil para el lector conocer cómo estos fenómenos se comparan con las "formas" estudiadas por Hegel en la *Fenomenología del espíritu*. Los primeros tres capítulos de ese trabajo se ocupan de las contradicciones teóricas o epistémicas en el nivel de la conciencia individual; para ser precisos, Hegel sigue las contradicciones sucesivas que surgen en los intentos filosóficos por explicar el sentido del mundo exterior. El capítulo IV de la *Fenomenología* avanza en la exploración de las contradicciones de la mente que surgen del nivel práctico de los deseos y la conducta; aquí es un momento fundamental la dialéctica amo-esclavo. En el capítulo V (o más bien en las subsecciones dos y tres de ese capítulo) Hegel hace la transición a las contradicciones sociales.

En este sentido muy general, los capítulos destinados en este trabajo a las contradicciones corresponde a lo que dice Hegel en la *Fenomenología del espíritu*. Sin embargo, los eruditos en Hegel verán de inmediato que no ofrezco ni una exégesis ni una paráfrasis de ese libro. Incluso cuando mi punto de partida explícito es el texto hegeliano, como en la discusión de la dialéctica amo-esclavo, me esfuerzo más bien por interesar al lector no hegeliano que por permanecer fiel a las intenciones del original. Podría haber intentado ambas cosas, pero no lo hice. Mi nivel de ambición es más bajo que el de Hegel, especialmente porque no presento una teoría de la *relación* entre las diversas clases de contradicciones. Para Hegel, el *orden* en que se presentan las diversas contradicciones es un rasgo decisivo de la exposición, como lo son también las transiciones entre un tipo de contradicción y el siguiente. En este análisis no se trabaja con dicha intención sistemática. Ofrezco al lector un racimo de ideas no totalmente conectadas,

no un sistema filosófico. La consistencia del proceso, si la tiene, debe ser buscada en el nivel metodológico, más bien que en el teórico. O sea, pienso que la *articulación entre las contradicciones lógicas y el estudio del cambio* es importante y fértil y tiene que ser estudiada en diversas variantes, incluso si no están ligadas en una teoría sustantiva .

Si el hegeliano erudito quedara insatisfecho después de una lectura de estos capítulos, los científicos sociales pueden también tener algunos problemas para entender exactamente qué se les ofrece. Incluso, no se presenta una teoría precisa para cada variedad o subvariedad de contradicciones: a lo sumo, se da algún tipo de contexto explicativo o "prospectivo", para utilizar un término que puede cubrir muchos pecados. En la última sección del capítulo 5 se intenta especificar algunas de las nociones mediadoras entre contradicciones sociales y cambios sociales, pero incluso aquí el lector con intereses empíricos (el destinatario principal de este libro) pensará que todo es muy esquemático; ese sentimiento se hará abrumador en este capítulo en el que no se dice nada acerca del mecanismo preciso del cambio. No ofrezco ninguna disculpa por esto. Más bien pediré excusas si este trabajo resulta ser un mal libro dentro de su clase, pero no me siento arrepentido de haber escrito un libro de este tipo. Si ningún científico social se siente tentado de llenar algunos de los vacíos de mis bosquejos y esquemas, esto será, en verdad, una seria acusación, pero me declararía no culpable ante la acusación de que es una pérdida temporales hacer bosquejos y esquemas.

La teoría de las contradicciones de Hegel

No presentaré el punto de vista de Hegel sobre las contradicciones como él lo plantea con todo su aparato conceptual en la *Ciencia de la lógica*, porque creo que la intuición básica de Hegel es mucho más simple de lo que podría creerse por esa exposición. Fundamentalmente, hace poco más que reiterar las paradojas eleáticas, con una conclusión levemente diferente. Desde el punto de vista de Hegel, la más importante de estas paradojas es la Paradoja de la Flecha.[5] Esta significa probar que una flecha está y no está en reposo a cada momento de su vuelo. *Está en reposo* porque, por definición, ningún objeto puede moverse durante un instante, ya que el movimiento requiere duración; es imposible ir de un lugar a otro en un intervalo nulo de duración. *No está en reposo* porque, si lo estuviera, sería imposible que llegara a algún lugar. El último argumento puede explicitarse de dos formas diferentes. Primero, podemos observar que si en *cualquier* instante de su vuelo la flecha no se moviera, es decir,[6] entonces estaría en reposo en *todos* los instantes siguientes, porque habría perdido su estado de movimiento. Segundo, y en forma menos anacrónica, podemos argumentar que si la flecha estuviera en reposo en todos los instantes durante su vuelo, entonces nunca dejaría el arco ni atravesaría distancia alguna, ya que un

compuesto de inmovilidades sólo puede dar la inmovilidad como resultado. El primer argumento apela a la primera ley de Newton; el segundo, a la idea intuitiva de que sumar ceros, no importa cuántos, no puede llevarnos más allá de cero.[7]

El lector se equivocaría si descartara esto como mera sofisticación. Las paradojas de Zenón —la de la Flecha no es la menos importante— apuntan hacia problemas profundos y difíciles en el análisis conceptual del tiempo y el espacio. Sería perverso decir que constituyen una fundación sólida de la dialéctica hegeliana, pero debe admitirse que la teoría hegeliana de la contradicción estaba propuesta como una solución a lo que está lejos de ser un seudo problema. Esquemáticamente, tres elementos diferentes entran en la Paradoja de la Flecha: la posibilidad de un movimiento real, la posibilidad de contradicciones reales y la validez de la paradoja. El punto de vista moderno es que ya que el movimiento es posible y las contradicciones son imposibles debe haber algo equivocado en el argumento de la paradoja, aunque no hay consenso general acerca de qué es lo que está equivocado. El punto de vista eleático era que, ya que las contradicciones reales son imposibles y el argumento es válido, tenemos una prueba por *reductio ad absurdum* contra la posibilidad de movimiento. La posición de Hegel era aún más radical: argumentando que el movimiento es real y el argumento válido, concluyó a favor de la existencia de contradicciones reales.[8] Es altamente significativo que esta admisión de "contradicciones reales" esté ligada al análisis del cambio. Hegel parece preguntarse ¿cómo puede algo cambiar si es sólo lo que es? ¿Cómo, por así decirlo, puede salirse de su propia piel si en su propia naturaleza no contiene su opuesto? Es también muy importante observar que Hegel nunca discute la Paradoja de la Flecha, comparada por algunos autores[9] por ejemplo con la paradoja de la semilla de mijo: como una sola semilla no hace ningún ruido al caer, 1.000 semillas al caer tampoco deberían hacer ningún ruido, porque cero multiplicado por 1.000 sigue siendo cero. En esta paradoja no hay un problema de cambio sino sólo una comparación entre casos sucesivos de semillas que caen al piso en diversas cantidades. Hegel nunca dice que al mismo tiempo escuchamos y no escuchamos una semilla al caer.

Dados el comentario de Hegel acerca de la Paradoja de la Flecha y más en general la sección sobre las contradicciones en la *Ciencia de la Lógica*, pienso que se le pueden razonablemente adjudicar las siguientes afirmaciones:

Primera tesis: Hay contradicciones en la realidad.
Segunda tesis: Una descripción adecuada de la realidad tiene que contener proposiciones autocontradictorias.

Se pueden encontrar afirmaciones similares en autores posteriores,[10] aunque quizá no con la misma inescrutable profundidad de intención. Voy ahora a lo que tomo como la crítica analítica corriente a estas dos tesis. Una crítica es la ejemplificada por el muy conocido artículo de Karl Popper, "Qué

es la dialéctica".[11] Esquemáticamente, argumenta que la primera tesis es o bien carente de significado o trivial, mientras que la segunda es falsa. Si por "contradicción" queremos en verdad *significar* contradicción, la primera tesis carece de significado y descansa sobre un error categorial similar a, digamos, el error de preguntar qué corresponde al signo de negación en la realidad. También podría decirse que, en sentido literal, la idea de "contradicciones reales" es un caso de la falacia patética, que adscribe a la realidad rasgos que caracterizan a nuestro pensamiento o a nuestra actitud hacia ella. J. Pen[12] ha observado con razón que es equivocado decir que una situación con monopolio bilateral es "indeterminada": la *situación* está perfectamente determinada en el sentido de que todas las variables tienen que tener algún valor en cada momento del tiempo, pero nuestra *teoría* puede ser indeterminada en el sentido de no estar capacitada para predecir estos valores. Así como la realidad no tiene espacios en blanco, tampoco tiene espacios doblemente llenos: a partir del hecho de que la teoría adscribe rasgos contradictorios al mundo real podemos extraer algunas conclusiones acerca de la teoría, pero no acerca de la realidad.

Por otra parte, si por "contradicciones" queremos simplemente significar conflicto, divergencia, oposición o lucha, la primera tesis se hace totalmente general y trivial. Nadie podría negar los hechos, digamos, de los conflictos de clase o de la lucha por la supervivencia. La primera tesis podría quizá ser reforzada para decir que en la realidad hay contradicciones *por doquier*, es decir, que *todos* los fenómenos de la naturaleza, la sociedad o el pensamiento contienen luchas y oposiciones. En ese caso tendríamos que pedir una definición exacta de la característica que se dice omnipresente. O bien el concepto está definido con precisión, de manera que la primera tesis se haga susceptible al análisis empírico y a la falsación, o el significado queda tan vago que luego del hecho siempre se puede apuntar hacia una característica que sea "contradictoria" en *algún* sentido. Un sostenedor de la primera tesis (en cuya lectura afirma que todos los fenómenos contienen luchas u oposiciones) puede quizá replicar que la búsqueda de una definición precisa no es lo más adecuado, y que la noción de contradicción se parece más bien a la noción de un juego, no habiendo características comunes a todas las contradicciones, sino tan sólo un parecido familiar.[13] Una respuesta tal puede ser comprendida de dos formas: como una descripción de cómo se utiliza la palabra "contradicción" en el lenguaje cotidiano, o como una vía para hacer científicamente útil el concepto. La observación sugerida en la primera interpretación de la respuesta es suficientemente correcta, pero no es de gran ayuda en nuestra empresa presente. En la segunda interpretación, la respuesta debe ser comprendida como una propuesta para que el concepto de contradicción sea definido disyuntivamente: nada que satisfaga por lo menos k entre n criterios dados es contradictorio. Tales definiciones disyuntivas pueden en verdad ser útiles en la ciencia[14] a condición de que una lista de criterios esté disponible. Como quienes abogan por el razonamiento dialéctico nunca dicen qué criterios tienen en mente, es difícil tomarlos seriamente. En efecto, sospe-

cho que la definición disyuntiva que gobierna la utilización del término sería más bien la siguiente: cualquier cosa puede llamarse contradictoria si tiene por lo menos una característica en común con por lo menos uno de los fenómenos que ya han sido clasificados como contradictorios. No es sorprendente que este concepto imperialista y expansionista termine devorándose toda la realidad; menos sorprendente que por lo menos un tercio de la humanidad tenga que adoptar este sinsentido como una norma de pensamiento.

En cuanto a la segunda tesis, que las descripciones correctas del mundo pueden contener afirmaciones autocontradictorias, claramente requiere el sinsentido de la primera tesis y por lo tanto podría ser descartada sin mayor escándalo. Sin embargo, Popper ofrece también un argumento separado contra la tesis, refiriéndose a la verdad lógica de que, dado un conjunto de proposiciones del cual pueden derivarse dos afirmaciones contradictorias, toda proposición, cualquiera que sea, también se sigue. El razonamiento dialéctico es en verdad fértil, pero no el tipo de fertilidad que es un *desideratum* de la teoría científica; más bien, la fertilidad que Poincaré tenía en mente cuando hizo la famosa broma de que la fertilidad de la lógica formal podía probarse por las antinomias que ella misma generaba. Me parece que el punto de Popper es que no se puede ser selectivamente autocontradictorio: la más pequeña fisura en una teoría la hará derrumbarse toda.[15] Añade una observación relacionada con una anotación posterior, al efecto de que las contradicciones pueden tener alguna utilidad real al inducirnos a cambiar la teoría, cuando se hacen evidentes sus inconsistencias. Aquí, entonces, tenemos una mirada a la posibilidad de otra articulación menos confusa entre contradicciones y cambio: en vez de argumentar que el cambio es contradictorio y tiene que ser descrito por afirmaciones autocontradictorias, tenemos mayor posibilidad de decir que el descubrimiento de que una teoría contiene contradicciones generará un cambio en la teoría.

Incluso si acepto la crítica de Popper sin reservas, pienso que en cierto sentido es algo trivial. Es verdad que Hegel parece haber mantenido el punto de vista absurdo refutado por Popper, pero hay en Hegel más que eso. El sentido en el cual pienso que es posible hablar sin ambigüedad acerca de contradicciones reales está dado por la siguiente tesis:

Tercera tesis: Hay situaciones en la realidad que sólo pueden ser descritas por medio del concepto de contradicción lógica.

Quien haya aceptado la primera y segunda tesis aceptaría sin duda la tercera, pero demostraré que lo inverso no es verdadero. Más aún, argumentaré que las "contradicciones reales" en este sentido – es decir, las situaciones cubiertas por la tercera tesis– están cercanamente articuladas al proceso de cambio; en parte porque son en sí mismas una variedad de cambio y en parte porque generan un movimiento de cambio en dirección a la consistencia. Se darán ejemplos, también, en este capítulo y en el 5 para corroborar la afirmación de que algunos de los escritos de Hegel y Marx

pueden ser analizados en términos de esta tesis. Puede quizá ser pertinente una última afirmación de intención. Mi primer deseo es que el análisis resulte interesante por derecho propio también para la mayoría de los lectores con un limitado interés por la lógica formal y poco interés por Hegel o la dialéctica. En segundo lugar, los análisis tendrían que tomarse en función de la totalidad del libro, por caso, como ejemplo de la utilización del análisis lógico en las ciencias sociales. Si, en tercer lugar, puedo ofrecer algo a esa pequeña pero apasionada minoría que se ocupa profundamente de Hegel y la dialéctica, este no será un resultado despreciado. Desde luego, no creo que he encontrado *el* uso correcto del término "contradicciones reales" ni desentrañado el secreto oculto de la dialéctica; no existen esos animales para encontrarlos o desentrañarlos. Tampoco afirmo que mi análisis agota cuanto hay de interés en Hegel y Marx. Sin embargo, *sí* espero que el análisis tenga en precisión lo que carece en alcance y fidelidad exegética, de manera que el crítico esté en capacidad –y obligado– a decirme *exactamente* lo que está equivocado en mi explicación de las contradicciones.

Deseos contradictorios: el amo y el esclavo

Las contradicciones de la mente, a las que me referí anteriormente, tienen una estructura muy sencilla. Básicamente es ésta. Supongamos que el conjunto de afirmaciones (*p,q,r,s...*) todas juntas denotan una contradicción. Entonces, una de las dos situaciones siguientes se ubicará bajo el encabezamieno de la tercera tesis ya enunciada:

Deseos contradictorios: A desea (espera, lucha por) que *p,q,r,s...* sean verdaderas simultáneamente.[16]
Creencias contradictorias: A cree que *p,q, r, s...*son verdaderas.

En las primeras dos secciones me ocuparé de los deseos contradictorios, reservando las creencias contradictorias para las últimas secciones. Pueden hacerse al inicio algunas observaciones generales que son válidas tanto para las contradicciones afectivas como para las cognitivas. Primero, adviértase la diferencia entre las siguientes situaciones:

(1) A cree/desea (*p,q,r,s...*). A partir de (*p,q,r,s...*) se puede deducir una contradicción.
(2) A cree/desea (*p,q,r,s...*) *w* es el caso. De (*p,q,r,s,w...*) se puede deducir una contradicción.

Si se toma literalmente, la tercera tesis anteriormente planteada, cubre ambas situaciones, pero me limitaré a poner mi atención en la primera. Sin duda, tenemos muchas *falsas* creencias y albergamos deseos que *de hecho* no pueden realizarse, pero estos fenómenos son muy generales para ser de interés. En cambio, intentaré demostrar la importancia de

las creencias *contradictorias* y de los deseos que *como una cuestión de lógica* deben permanecer irrealizados.[17] Se darán ejemplos para demostrar que esta diferenciación tal vez no es rígida, pero para mi actual propósito pienso que la realidad de la diferenciación en los casos corrientes es más importante que su posible colapso en las situaciones enfocadas por Quine y otros.[18] En una segunda observación, podemos diferenciar entre "grados de contradictoriedad". Compárense, por ejemplo, los casos siguientes:

(3) A cree/desea p; A cree/ desea $(p \supset \bar{q})$; A cree/desea q.
(4) A cree/desea p; A cree/desea $\bar{p}$.
(5)A cree/desea $(p \,\&\, \bar{p})$.

El último caso puede presumiblemente llamarse patológico; sólo los esquizofrénicos y los hegelianos pueden tener lo contradictorio como un objeto explícito de creencia y deseo. Por otra parte, las situaciones como la (3) son probablemente bastante comunes. Un matemático puede, por ejemplo, contemplar diversas ideas que luego se comprueban como lógicamente incompatibles. Sospecho que situaciones como la (4) son más comunes de lo que puede pensarse. Es posible creer que p es parte del campo de experimentación y $\bar{p}$, parte de un campo totalmente diferente, sin que la contradicción sea descubierta porque no hay interacción entre los campos.

Tener una torta y comérsela es una cuestión a veces imposible de lograr, pero siempre posible de emprender. El principal ejemplo que va a utilizarse para demostrar ese punto es la dialéctica del amo-esclavo de Hegel en la *Fenomenología del espíritu*, juntamente con algunas afirmaciones anteriores y posteriores acerca de lo que básicamente es la misma visión. Permítasenos decir que la "fuerza dinámica" del trabajo de Hegel es el surgimiento y luego la superación de las contradicciones en situaciones como la (1) y la (3) antes mencionadas. Hegel insiste mucho en que él no intenta refutar a la conciencia, cuya odisea sigue. No está aplicando estándares externos de verdad ni hechos que no sean propios de la conciencia; más bien, intenta demostrar cómo el mismo sujeto descubre la inconsistencia de su "forma" (*Gestalt*) corriente al comparar su propia estructura mental con su criterio de adecuación.[19] El que"aflore" la contradicción para transponerse en un objeto explícito de la conciencia induce un cambio en la forma de la conciencia, pero la nueva forma siempre alberga una contradicción suya y así, hasta alcanzar el nivel de conocimiento absoluto. En el capítulo 5, vuelvo a la lógica de este proceso de aprendizaje jerárquico, y sugiero que Hegel anticipó asombrosamente análisis posteriores de un aprendizaje inherentemente no exitoso. Sin embargo, sólo me ocuparé de una contradicción mental específica, la relación amo-esclavo.

Para obviar un frecuente malentendido, comencemos señalando que de acuerdo con Hegel no hay contradicción *entre* el amo y el esclavo, sino

sólo una contradicción *en* la mente del amo. (También hay una contradicción inherente a la conciencia del esclavo, pero no exploraré este aspecto aquí.) En todo caso, el punto de partida es una lucha entre dos conciencias , una pelea por el reconocimiento, en la que una termina sometiéndose, prefiriendo una vida sometida a la muerte en el combate. El esclavo se ha identificado a sí mismo con la materia bruta mientras que el amo –jugándose la vida en la lucha– ha trascendido la existencia natural y demostrado su libertad absoluta. La cuestión es que el amo es doblemente poderoso. Por medio del esclavo tiene poder sobre los objetos materiales de consumo, los que puede disfrutar sin haber trabajado para adquirirlos. A través de los objetos materiales, el amo disfruta el poder sobre el esclavo, ya que con el consumo arbitrario y conspicuo de esos objetos demuestra su poder sobre el ser que los ha producido. Sin embargo, el amo está atrapado, porque no puede obtener satisfacción real de su poder sobre un ser al que trata como una cosa, un *instrumentum vocale*, como los romanos llamaban a sus esclavos. En verdad, la idea central de esta sección es que la noción de un *reconocimiento unilateral*, que es lo que el amo desea del esclavo, es contradictoria; esto puede verse más fácilmente, quizá, si pensamos en la ridícula idea de una nación que busca reconocimiento diplomático de una de sus colonias. No puede forzarse a alguien a que respete a otro, aunque se puede intentar. El amo querría tener tanto la satisfacción que se deriva del reconocimiento que le debe el esclavo como la satisfacción que se deriva de su poder absoluto sobre el esclavo. Pero realmente esto es como intentar tener la torta y comérsela; el amo desaparece de la historia como un callejón sin salida y el esclavo se transforma (por un momento) en el portador del desarrollo posterior.[20]

En lo que resta de esta sección exploraremos algunas variantes de esta dialéctica, tomando elementos de la poesía y la ficción tanto como de la psicología o la filosofía. El trabajo de Hegel ha sido llamado una novela metafísica,[21] y a la inversa, los poetas metafísicos, como Donne (o Emily Dickinson), conocen tanto sobre las contradicciones de la mente como Sartre o Bateson. A los escritores citados en la frase anterior debemos añadir el nombre de Eugene Genovese quien analizó la psicología de la esclavitud negra en términos de la relación amo-esclavo. Entre las muchas ideas elaboradas por Genovese está la de que, a través de esa relación, de alguna manera los amos se tornaron esclavos de sus esclavos: "Los amos necesitaban desesperadamente la gratitud de sus esclavos para definirse ellos mismos como seres humanos morales. Los esclavos, al rehusarse, infligieron una puñalada en la autoimagen del amo."[22] Esta visión paradójica no fue sin duda original de Hegel. Encontró una expresión perfecta en el poema de Donne:

> Cuidado con odiarme,
> O triunfarás demasiado en la victoria.
> No porque yo vaya a ser mi propio vengador,
> Y al odio con odio nuevamente pague.

Porque perderás el estilo de conquistador,
Si yo, tu conquista, perezco por tu odio.
Entonces, para que mi existencia en nada te disminuya,
Si me odias, cuidado con odiarme.[23]

Rousseau expresó la idea fecunda de que *Quiconque est maître ne peut être libre*, y en verdad esta idea parece haber sido una paradoja común en el siglo XVIII.[24] Desde otro punto de vista, la noción de Hegel puede compararse con la idea de que el poder más eficaz puede ser ejercido invisiblemente, a través de normas interiorizadas más bien que por medio de sanciones manifiestas, y que el uso del poder es un acto de consumo más que un acto de inversión.[25] Pese a esos precedentes, debe darse crédito a Hegel por haber señalado la naturaleza contradictoria e inherentemente irrealizable del deseo del amo y por haber visto esta contradicción como una de las muchas confusiones que nos dejan atrapados en nuestro intento por encontrar una *satisfacción durable* (lo que según Leibniz sería una *contradictio in adjecto*).[26]

El lector puede haberse preguntado en qué sentido, exactamente, el deseo del amo es *contradictorio*. ¿Tenemos realmente una situación del tipo (1) ya mencionado, o tenemos que agregar –como en las situaciones del tipo (2)– algunas premisas fácticas adicionales para generar la contradicción? Sin duda, a partir de la idea del reconocimiento unilateral no podemos deducir inmediatamente algo como "$p \,\&\, \bar{p}$". Aquí estamos manejando algún tipo de *imposibilidad conceptual*; un caso en el que la lectura del contenido implícito y las presuposiciones del concepto generan una contradicción, incluso si la forma manifiesta evidente es consistente. En el caso presente esta lectura puede hacerse aproximadamente como sigue. El concepto de reconocimiento difiere de nociones relacionadas tales como aceptación o apreciación, en que lo que es reconocido no es la existencia o algún aspecto estético de la persona, sino su valor moral. Al buscar el reconocimiento, implícitamente se reconoce el valor moral de la persona para quien deviene el reconocimiento, pero al buscar la apreciación no se aprecia implícitamente el valor estético de la persona que apreció al otro. Esto es así porque la competencia moral *es* valor moral. Reconocer a alguien como competente desde el punto de vista moral –como implícitamente lo hago cuando busco su reconocimiento de mi valor moral– es reconocerlo como moralmente apropiado. La competencia estética, de otro lado, de ninguna manera implica que la persona competente tiene un valor estético.

Por estos motivos, la apreciación unilateral es una noción consistente, mientras que el reconocimiento unilateral no lo es. Considerando las presuposiciones implícitas de un concepto puede ser factible generar una contradicción lógica que se derive de la forma manifiesta del concepto. Obviamente, el problema es cómo distinguir estas "presuposiciones implícitas" de las premisas puramente fácticas que entran en situaciones del tipo (2) ya mencionadas. ¿Es un hecho *empírico* (aunque muy importante) acerca de los seres humanos que la competencia moral es inseparable del

valor moral o tendríamos que decir que esta articulación es constitutiva de lo humano, como el hecho de que somos mortales? Dichas cuestiones demuestran lo que yo pensaba cuando afirmaba que la diferenciación entre la imposibilidad lógica y la falsedad empírica puede no ser rígida. Podemos ir de nuevo a Donne para iluminación e ilustración. En *La paradoja* leemos estos enigmáticos versos:

No podría decir que amé, pues ¿quién podría decir
Que fue asesinado ayer?

Si vamos a *El corazón roto*, la paradoja es establecida en términos más fácilmente comprensibles:

Está totalmente loco quien dice
Que ha estado enamorado una hora,
Pero que el amor no declina tan pronto,
Sino que puede devorar a diez, en menos espacio.
¿Quién me creerá si juro
Que he tenido la plaga por un año?
¿Quién no reiría de mí si dijera que
vi una redoma de pólvora arder por un día?

La cuestión es que algunos verbos no pueden ser utilizados en primera persona del singular y en el tiempo pasado sin dar lugar a lo que Jaakko Hintikka[27] ha llamado una *inconsistencia existencial*. Decir "yo fui asesinado ayer" es realmente inconsistente por razones puramente lógicas; decir "vi una redoma de pólvora arder por un día" es inconsistente sólo si añadimos alguna premisa fáctica, como la probabilidad de sobrevivir a este hecho si uno estuvo suficientemente cerca para observarlo; decir "yo amé" es existencialmente inconsistente si añadimos la premisa de que el amor transforma de manera tal a una persona como para hacerlo otro individuo, pero ¿qué tipo de premisa es ésta?

Después de esta digresión sobre las ambigüedades de la noción de una contradicción mental, volvamos a la sustancia del problema. La breve referencia anterior al trabajo de Eugene Genovese no estaba dirigida a una discusión exhaustiva de su versión de la dialéctica del amo-esclavo. Queda mucho por decir acerca de cómo se transforman las presunciones metafísicas en problemas empíricos, haciendo así el tema de las interpretaciones "dialécticas versus positivistas" de la esclavitud un tema factible de ser estudiado. En el capítulo 6 vuelvo a una confrontación semejante, relacionada con el significado de "irracionalidad"en el comportamiento de los plantadores. Aquí discutiré un problema más directamente relacionado con lo que se dijo anteriormente sobre las contradicciones inmanentes de la conciencia del amo. En un estudio reciente del trato dado a los esclavos norteamericanos, G. Canarella y J. Tomaske[28] diferencian entre los amos sádicos y los paternalistas: los primeros incrementan su utilidad por medio

de una reducción del bienestar de los esclavos y los segundos extraen utilidad positiva del bienestar de los esclavos. Esto puede contrastarse con la descripción de Genovese del plantador típico: "Eran fuertes, orgullosos y arrogantes; de espíritu liberal en todo aquello que no tocara su honor; amables y corteses; generosos y benévolos; fácilmente iracundos y extraordinariamente crueles; atentos al deber y descuidados de cualquier tiempo o esfuerzo que no sirviera a sus intereses directos."[29] En las líneas hegelianas bosquejadas antes y seguidas por Genovese, la diferenciación pertinente no sería entre dos tipos de amo, sino entre dos instancias de comportamiento de cada amo, oscilando entre la benevolencia y la crueldad, de acuerdo con el lado de la contradicción que gana la mano. En términos que aceptamos como débiles, aunque muy sugerentes, podríamos decir que la visión "positivista" supone que los individuos son (i) consistentes (ii) diferentes entre sí, mientras que una interpretación "dialéctica" insiste en que cada individuo es *diferente de sí mismo* y por lo tanto inconsistente. El mérito de Genovese es haber demostrado que esta oposición no debe resolverse en el nivel *a priori* del razonamiento metafísico, sino que, al contrario, debería ser perfectamente accesible como problema de investigación en la psicología social.[30]

El ser y la nada, de Jean Paul Sartre, es en muchos sentidos una importante variación sobre estos temas hegelianos. En verdad, Sartre afirma que el hombre tiene "una cierta capacidad para formar conceptos contradictorios";[31] aunque siempre es tentador aplicar el epíteto de "contradictorios" a los análisis vertiginosos que ofrece Sartre, pienso que muchos de ellos tienen mucho sentido, una vez que comprendemos que suscribe no la segunda sino la tercera tesis sobre las contradicciones, que discutimos anteriormente. Como un ejemplo típico del tipo de razonamiento encontrado en el trabajo de Sartre, podemos tomar su discusión acerca del amor. El amante no desea la sumisión total de la persona amada, tampoco un compromiso libre y voluntario; desea lo imposible: "poseer una libertad".[32] Sartre refleja fielmente el nuevo análisis psicológico que ha estado entre nosotros desde Stendhal. A una pregunta frecuentemente citada de *Andromaca* de Racine, *Je t'aimais inconstant, qu'aurais-je fait fidele?* (Te amé mientras eras inconstante, qué no hubiera hecho si hubieras sido fiel?), Stendhal, Proust y Sartre dan respuestas opuestas a la que implicaba Racine: la reciprocidad es la muerte del amor ...y el objeto del amor. El amor es un proyecto imposible, "una pasión inútil", para utilizar la expresión que, de acuerdo con Sartre, resume las condiciones humanas en su conjunto.[33] El amor requiere "la síntesis imposible de la asimilación y el mantener la integridad del objeto asimilado".[34] (Para aquellos interesados en las diferencias nacionales del estilo intelectual, podemos observar que para los franceses esto es "tener la torta y comérsela".) De manera más general, la existencia del hombre, de acuerdo con Sartre, se define por un deseo contradictorio de ser simultáneamente una cosa (*en soi*) y una conciencia (*pour soi*); ser tan sólido y positivo como un enano blanco donde no queda lugar entre los átomos *y* saber esto. Me parece que no estaría lejos

de la verdad si dijera que para Sartre el dilema del hombre es encontrarse perseguido por el deseo de escuchar su propio discurso funerario. Si, como Tom Sawyer y Huckleberry Finn, pudiéramos estar secretamente presentes en nuestro propio funeral, podríamos tener tanto el resumen final de nuestras esencias (*tel qu'en lui-même enfin l'éternité le change*) *y* mantener la conciencia necesaria para disfrutarlo; descansar en nosotros mismos en libertad total de la libertad y mantener la distancia de nosotros mismos, que es la señal de la conciencia. No es necesaria una explicación muy extensa para ver por qué es existencialmente inconsistente, en el sentido de Hintikka, afirmar que uno ha estado presente en su propio funeral y por qué el deseo de estar presente es inherentemente contradictorio.

El impacto de estas ideas en la psicología moderna ha sido (hasta donde yo sé) escaso, pero se pueden seguir algunos desarrollos paralelos. Utilizando el *Gesamtregister* para las *Obras completas* de Sigmund Freud,[35] es sencillo investigar lo que Freud dijo acerca de la negación y la contradicción. Se destacan esencialmente dos ideas. Primero, el inconsciente no contiene negaciones ni contradicciones: en éste, cada elemento es completo, positivo y no relacionado con todos los demás.[36] Segundo, el terapeuta sabe que durante el análisis la negación es sólo otra forma de confirmación; negar algo es sólo una forma de afirmarlo. Podría decirse que esta última idea contiene un germen de la teoría del doble vínculo, propuesta por Gregory Bateson.[37] La riqueza del pensamiento de Bateson es igualada tan sólo por su capacidad elusiva y ambigüedad; me abstendré, por lo tanto, de una exposición o exégesis y sólo afirmaré que a mi parecer Bateson está mucho más cercano a Hegel y Sartre que a la cibernética, la teoría de los tipos de Russell, la teoría de los juegos de Von Neumann y la moderna teoría de la evolución, todas abundante y no muy convincentemente citadas en su trabajo. Esencialmente, Bateson ha redescubierto, juntamente con Joseph Heller y otros, una visión paradójica de raíces muy atávicas, tal como lo hemos visto. Para mostrar que es realmente la misma visión, podemos considerar un ejemplo sencillo de cómo establecer un doble vínculo; una madre que le dice a su niño: "Recuerda que ni debes pensar en esta cuestión prohibida." Aquí, la presuposición implícita del mandato –para cumplirlo, el niño tiene que recordar (es decir, pensar) en el objeto prohibido– contradice el contenido manifiesto. En verdad.

El Corazón no puede olvidar
Excepto si contempla
Lo que rechaza[38]

Es inherentemente imposible realizar una *decisión de olvidar*, excepto, quizás, utilizando estrategias complicadas e indirectas tales como las que concibió Pascal en sus instrucciones acerca de cómo llevar a cabo la decisión igualmente paradójica de creer.[39] Podríamos distinguir aquí entre negación pasiva y activa, entre la negación y la indiferencia y afirmar como

un axioma que el último es el estado resultante del primero. Esta, de nuevo, es una idea recurrente en la *Fenomenología* de Hegel. La mente siempre está teñida de lo que rechaza; el rebelde, con la sociedad a la cual desea abolir; el rebelde desilusionado, con el Dios al cual falta. La relación con el concepto de la negación de la negación, mencionado al final del capítulo 3, tendría que ser obvia.

Deseos contradictorios: la economía de la irracionalidad

Tendemos a pensar que una persona que demuestra actitudes contradictorias es "irracional" en cierto sentido del término. ¿Hasta qué punto es también verdadero lo opuesto? O sea, cuando un economista habla de comportamiento irracional, ¿se refiere a algo que se presta a ser analizado en términos de una contradicción lógica? En esta sección intentaré responder esta pregunta para algunos de los diversos comportamientos llamados "irracionales" por los economistas. En el lenguaje de los economistas, el sentido principal y estándar de irracionalidad es el de un conjunto de preferencias intransitivas, por ejemplo, un individuo que prefiere *a* a *b*, *b* a *c* y *c* a *a*. Ahora, de esta intransitividad no puede deducirse ninguna contradicción, a no ser que la persona también tenga una metapreferencia por estructuras transitivas de preferencia más que por estructuras intransitivas. Varios estudios empíricos demuestran que las personas muestran a veces preferencias intransitivas y que cuando esto se les indica, usualmente modifican sus preferencias en el sentido de la consistencia. Puede decirse, presumiblemente, que este cambio revela su metapreferencia por la transitividad. Sin embargo, tenemos que añadir que la metapreferencia por la transitividad tiene que ser jerárquicamente superior a todas las demás preferencias, porque de otra manera una persona flagrantemente sorprendida en intransitividad podría responder algo similar a lo siguiente: "Sí, acuerdo con que la transitividad es una buena cosa, pero no a cualquier costo. Aquí, como en todo lugar, los intercambios son pertinentes y, si tengo un sentimiento lo suficientemente fuerte acerca de un conjunto dado de alternativas, estoy dispuesto a abandonar la transitividad."

Pienso que no hay contradicción lógica alguna en el concepto de esa persona, pero por razones explicadas en el próximo apartado, la idea es poco plausible. Sin embargo, se hace menos rebuscada si miramos a formadores de decisiones colectivos más que actores individuales, así sean empresas, organizaciones o Estados. Las siguientes dos proposiciones de la teoría de la elección social son bien conocidas. (a) Si la elección social es por votación mayoritaria, existen constelaciones de preferencias individuales que inducen preferencias de intransitividad social. (b) Si se impone un consenso moderado –lejos de la unanimidad– por encima de las preferencias individuales, entonces la intransitividad de las preferencias sociales no puede surgir.[40] Sin embargo, en este caso es muy sensato para la comuni-

dad en cuestión decir que la transitividad no tendría que ser un desiderátum predominante; que es más importante respetar las opiniones individuales que evadir las paradojas. Esperamos que un individuo cambie sus preferencias cuando ve que son intransitivas, pero la paradoja de una votación mayoritaria es que no da un motivo que compele a cambiar opiniones. Una comunidad bien puede hacer suyas las palabras de Whitman: "¿Me contradigo a mí mismo? Bien, lo hago. Soy extenso, contengo multitudes."

Hay una forma divertida e instructiva[41] de destacar la irracionalidad de la intransitividad. Una persona que prefiere *a* a *b*, *b* a *c* y *c* a *a* podría ser desangrada de la siguiente manera. Como prefiere *a* a *b*, debe estar dispuesto a pagar *cierta* cantidad, aunque sea muy pequeña, para poder intercambiar *b* por *a*; de la misma forma debe estar preparado para pagar algún dinero para poder intercambiar *a* por *c* y, por supuesto, una mayor cantidad para intercambiar *c* por *b*. Al final de este proceso se queda con *b*, como al inicio, pero faltándole algún dinero. Repitiendo este proceso, el individuo podría *mejorarse a sí mismo hasta la muerte* (a no ser que las preferencias fueran tales que las cantidades sucesivas de dinero entregado formen una serie que converge a una suma menor que la cantidad de dinero que tenía al inicio). Sugiero que esta noción de arruinarse uno mismo a través de una serie de pasos, cada uno de los cuales es visto como un avance, se puede generalizar.[42] Este proceso paradójico puede surgir de la inconsistencia, como en el ejemplo anterior, o de la inconstancia, como en el caso del cambio endógeno de preferencias. Si el lector se remite al relato en el capítulo 2 del análisis que Von Weiszäcker hace de este cambio, recordará que la anotación "*x P(y)z*" significa que el individuo en cuestión prefiere *x* a *z*, dado que su consumo en el período inmediatamente precedente era *y*. Von Weiszäcker ha probado que, dadas ciertas condiciones, que pueden ser interpretadas como una tendencia[43] para degradar el *status quo*, podemos construir una secuencia q_i de manera que $q_{i+1} P (q_i) q_i$ para toda *i*, en donde sin embargo la q_i converge a cero en cada componente. (En un ejemplo de dos componentes, la secuencia puede ser (1/2,3/2), (3/4,1/2), (1/4, 3/4), (3/8,1/4)..., en donde cada paso es aceptado porque implica un aumento del componente menor. Este es el caso de mejorarse uno mismo hasta la muerte como resultado de una preferencia miope por el cambio.

En sí mismo, a este caso de preferencias cambiantes endógenas no podría aplicar el epíteto de "contradictorio". Es sin duda trágico que una persona pueda ser llevada a la muerte por fuerzas causales que actúan sobre él sin su conocimiento o comprensión, pero, como se argumentará en el apéndice 2 del capítulo 5, la causalidad subintencional es un hecho del cual no podemos escapar. Sin embargo es natural preguntar si no tendríamos que hablar de una contradicción cuando el individuo está *consciente* de las implicaciones globales de sus acciones (es decir, no es sólo el vehículo pasivo de las preferencias cambiantes) y sin embargo actúa de acuerdo con las preferencias instantáneas. Cuando el problema se plantea de esta manera, es claro que es un caso especial (y algo artificial) del problema más

general de las preferencias puramente temporales. Por una diversidad de motivos,[44] las personas pueden preferir el presente o el futuro cercano al futuro distante. (También pueden como el avaro preferir el futuro al presente, pero haré abstracción de esto en lo que sigue.) Esta preferencia ha sido llamada "irracional" en términos generales: un año es un año es un año y no hay nada sino carencia de voluntad o imaginación para inducir una preferencia por un año sobre otro. Esto es muy simplista. Para un individuo, el hecho de ser mortal hace que sea en verdad racional adjudicarle más peso al presente que a un futuro en el cual ya él no esté presente. Para una sociedad que desea distribuir recursos entre el presente y el futuro dentro de la planificación de un horizonte infinito, dar igual peso a todas las generaciones puede dar el resultado paradójico de que cada generación destinara todo a la inversión y nada al consumo (por encima de la subsistencia).[45] Sin embargo, si nos restringimos al caso individual y abstraemos las consideraciones de mortalidad, lo cual está justificado por el hecho de que las preferencias temporales frecuentemente son más fuertes de lo que las estadísticas de mortalidad implican, nos queda un comportamiento que quizá puede ser llamado "irracional". ¿Tendría que llamarse también "contradictorio"?

Distinguiré entre el problema general de las preferencias temporales y el problema especial de la *inconsistencia de las preferencias temporales.*[46] El problema especial está implícito en la oración de Agustín: "Dame castidad y moderación, pero no todavía",[47] la cual interpreto de la manera siguiente. A partir de mañana seré "casto y moderado", lo cual implica entre otras cosas que no le dé gran peso al presente. Hoy, sin embargo, me gustaría satisfacer algunos de mis vicios, uno de los cuales es una preferencia por el presente. Por otra parte, soy suficientemente inteligente para ver que cuando el mañana llegue me sentiré aún igual. Por esta razón, ruego a Vos me otorgues las virtudes de las cuales carezco, a partir de mañana. Mi situación es similar a la de la persona que ha heredado una fortuna, de la cual quiere disponer como sigue. Durante el primer año quisiera hacer un gran banquete, gastando la mitad de su herencia, y luego quisiera hacer rendir lo demás de manera igual todo el resto de su vida. Sin embargo, si tiene suficiente claridad verá que cuando llegue el segundo año querrá gastar la mitad (o por lo menos una cantidad desproporcionada) de lo que le queda en el segundo año, dejando lo demás para distribuirlo de manera igual en los años siguientes, y así sucesivamente. Una contraestrategia sería comprar una anualidad con la mitad de lo que no piensa gastar en el primer año; esto sería equivalente a rogarle a Dios que me diera castidad en lugar de confiar en mi propia voluntad. Otra contraestrategia es lo que R.H. Strotz llama *estrategia de la planificación consistente*, la cual aproximadamente equivale a la elección de la mejor secuencia de consumo que satisfaga la restricción de presupuesto y las preferencias inconsistentes.[48]

Más bien que entrar en el detalle de estas contraestrategias, permítasenos mirar la inconsistencia para la cual son soluciones. Si escribimos $f_i(j)$ para el valor descontado en el tiempo i para una unidad de utilidad

experimentada en el tiempo *j*, podemos establecer dos exigencias a las preferencias temporales:

Preferencias temporales constantes: $f_i(i):f_i(i+1)=f_j(j):f_j(j+1)$
Preferencias temporales consistentes: $f_i(j):f_i(j+1)=f_j(j):f_j(j+1)$

A partir de esto obtenemos de inmediato la condición de

Preferencias temporales exponenciales: $f_i(i):f_i(i+1)=f_i(j):f_i(j+1)$.

La presunción de preferencias temporales constantes puede ser cuestionada. En la discusión anterior de las preferencias cambiantes endógenas, explícitamente presumí que lo que cambió fueron las preferencias sobre el conjunto de vectores de los bienes *en cualquier momento dado*. No sólo no había ninguna cuestión de cambio endógeno de preferencias temporales, sino que incluso apelé a las preferencias temporales constante para una explicación de la tendencia "irracional" a la búsqueda de ganancia a corto plazo. No conozco ninguna discusión en la literatura de las preferencias endógenas de cambio temporales, pero puede haber por lo menos dos posibles fuentes para dicho fenómeno. En primer lugar, las preferencias temporales que surgen del análisis de los índices de mortalidad cambiarán necesariamente en la medida en que avanzamos en edad sin morir; en segundo lugar, la posesión de preferencias temporales inconsistentes en el sentido de lo que se está discutiendo puede en sí misma inducir una tendencia al cambio de acuerdo con mi tesis general de que las contradicciones inducen cambios. En cuanto a la presunción de preferencias temporales consistentes, he intentado captar la idea de que una persona consistente no tendría que revisar sus planes constantemente, o mejor: la estructura interna de sus planes no tendría que ser tal que tuviera que ser *necesariamente* revisada incluso sin "cambio real" en el carácter de la persona. Hay aquí grados de racionalidad. En verdad, no es irracional cambiar sus planes porque usted se convence por experiencia, persuasión o argumentación de que sus preferencias anteriores eran inmorales o superficiales. Tampoco es formalmente irracional, aunque sí potencialmente desastroso, revisar sus planes por preferencias endógenas de cambio. Sin embargo, parecería como no ambiguamente irracional tener planes inherentemente irrealizables. Una condición necesaria y suficiente para que el plan sea realizable es que el peso relativo de dos años dados en el futuro no cambie con el avance del tiempo, porque sobre esta condición será en realidad llevada a cabo la distribución planificada del consumo en esos dos años.

Hacer planes "inherentemente irrealizables" parecería implicar que una contradicción está presente en algún lugar, pero en este caso la forma lógica de la contradicción no está clara de inmediato. Pospondré una mayor discusión para el capítulo 5, porque creo que las preferencias temporales inconsistentes tienen una estructura más parecida a las contradicciones de la sociedad que a las de la mente. Permítasenos afirmar brevemente que

cada encarnación sucesiva de los individuos desea estar libre para atar a sus sucesores mientras que no se siente atada por las decisiones de sus predecesores. Mientras que cualquier encarnación puede realizar este deseo, perpetramos la falacia de composición si concluimos que todas las encarnaciones pueden lograrlo. Pospongo también para el próximo capítulo una mayor discusión de cómo surgen las preferencias temporales no exponenciales.

Paso ahora al problema general de las preferencias temporales, que propongo subsumir bajo el encabezamiento más general de *akrasia* o debilidad de la voluntad. En lo fundamental, sigo el análisis propuesto por Donald Davidson.[49] Se ha reconocido desde hace tiempo que la debilidad de la voluntad tiene algo que ver con las contradicciones, no en el sentido de la tesis tercera anteriormente propuesta, sino más bien en el de la primera. Se ha preguntado, no si la estructura mental de los *akrates* descansa sobre actitudes o creencias contradictorias, sino si el concepto mismo de una persona tal es significativo. ¿Es posible (lógicamente) que una persona actúe contra su propio juicio, que conozca el bien y sin embargo haga el mal? ¿No es el hecho mismo de hacer el mal prueba suficiente de que uno no conoce el bien, o no piensa que es el bien? La manera de resolver el problema de Davidson –y de probar la posibilidad de *akrasia*– está basada en su teoría general de que las razones son causas.[50] Supóngase que tiene una elección libre entre x y y. Sobre la base de un conjunto de razones s prefiere x a y, pero sobre la base de las razones r (que incluyen s) prefiere y a x. Sin embargo hace x por las razones s. Usted actúa libre e intencionalmente (porque tiene una razón para actuar como lo hace), pero no toma en cuenta todas las razones relevantes. ¿Cómo puede dejar de lado todas las razones adicionales que entran en r? De acuerdo con Davidson,

> no hay ninguna paradoja en suponer que una persona algunas veces sostiene que todo lo que cree y valora sustenta un cierto curso de acción, cuando al mismo tiempo esas mismas creencias y valores lo llevan a rechazar ese curso de acción. Si r es la razón para que alguien sostenga que p, entonces el que sostenga que r tiene que ser, me parece, una causa para sostener que p, pero, y esto es lo crucial aquí, el que sostenga que r puede causar el que sostenga que p sin ser su razón; en verdad, el agente puede incluso pensar que r es una razón para rechazar p.[51]

Si utilizamos el término "motivación" para la relación entre una razón y una acción para la cual es la razón, entonces la razón r puede motivarme a hacer y más bien que x y llevarme a hacer x más bien que y, en cuyo caso mi razón para hacer x es también causa parcial de que haga x, ya que s está incluida en r. Esto es tan sólo una de las muchas curiosas formas en las que las vidas causal e intencional de la mente se pueden estropear una a la otra. Algunas observaciones más sobre este difícil tópico se ven en el apéndice 2 del capítulo 5. Un resumen preliminar puede ser el siguiente. Deberíamos diferenciar entre intenciones contradictorias, de un lado, y conflictos entre

los procesos causales e intencionales, de otro lado. Las preferencias inconsistentes temporales y (con algunas presunciones adicionales) las preferencias intransitivas son casos claros de deseos contradictorios o intenciones inherentemente irrealizables. Las preferencias temporales consistentes o, más generalmente, la debilidad de la voluntad deberían ser más bien vistas como irracionalidad intencional, en los términos utilizados por Davidson; aquí el actor está atrapado por fuerzas causales que lo eluden y hacen que todos sus planes lleguen a nada. El fenómeno de los cambios endógenos de preferencias tiene un carácter similar.

Creencias contradictorias: Hintikka

Habiendo argumentado que el hegelianismo y la teoría económica del comportamiento racional son dos fuentes fundamentales para el análisis de las actitudes contradictorias, avanzo ahora hacia los conocimientos contradictorios. El análisis lógico de lo que debe *significarse* con el término "contradicción cognitiva" entra de manera extensa en el trabajo de Jaakko Hintikka acerca de la lógica epistémica, mientras que el análisis de lo que ocurre cuando las personas albergan creencias contradictorias se refiere principalmente a la teoría de la disonancia cognitiva de Leon Festinger. No es evidente para todos los escritores que *pueden* existir sistemas contradictorios de creencias. Tómese por ejemplo, el siguiente pasaje de un influyente artículo de Alastair MacIntyre:

> Calvino estaba comprometido con las siguientes proposiciones: 1. Dios ordena las buenas obras; 2. Es de la mayor importancia posible hacer lo que Dios manda; 3. Las obras de Dios son irrelevantes para lo que es de mayor importancia para usted y su salvación o condena. Es un requerimiento de la lógica, no de las presiones psicológicas, que una de estas proposiciones sea modificada; la alternativa es la contradicción.[52]

Incluso dada la cuestionable caracterización de estas tres proposiciones como contradictorias, no hay razón para excluir su cosostenibilidad *qua* objetos de creencia, incluso si es probablemente cierto que esto establecería presiones psicológicas. Concluir, a partir de la imposibilidad de que algunas creencias sean verdaderas simultáneamente, la imposibilidad de que alguien las tome en consideración simultáneamente, es, por cierto mala lógica,[53] a no ser que se adopte la interpretación psicológica de la ley de la contradicción.[54] Pero no creo que MacIntyre haría esto.

Voy a seguir muy de cerca la discusión de Hintikka en el capítulo 2 de *Conocimiento y creencia*, donde discute los "criterios de consistencia" para su lógica epistémica. Ahora bien, esta lógica pretende tratar tanto acerca del conocimiento como de la creencia, y los criterios de consistencia van a ser muy diferentes para estas dos "actitudes proposicionales". En un lenguaje ordinario, sería contradictorio decir de alguien que "El sabe que p y que no p"; tal persona no tendría una estructura mental contradictoria,

simplemente no podría existir. (Esta es la diferencia entre las tesis primera y tercera sobre las contradiciones.) Por otra parte, como se explicó antes, no es contradictorio decir de alguien que "cree que p y que no p. Sin embargo, Hintikka no quiere admitir que afirmaciones como la última son consistentes. El quiere explorar lo que podría llamarse *creencia racional*, o en sus propios términos: *creencias defendibles*. Un sistema defendible de creencias tiene "inmunidad a cierto tipo de críticas" de manera que no puede ser refutado sobre bases puramente lógicas. (Para bosquejar una articulación con la teoría de Festinger, podemos observar que un sistema defendible de creencias no tendría que ser una fuente de disonancia cognitiva, a no ser que el individuo equivocadamente piense que una contradicción puede ser deducida de lo que él cree o que se adicionen criterios psicológicos de consistencia a los criterios lógicos.)

En el trabajo de Hintikka se dan dos conjuntos de interpretaciones a los operadores modales "M" y "N", correspondientes a conocimiento y creencia. En el análisis de conocimiento "$K_a p$" (que corresponde a necesidad) significa que "a sabe que p", mientras que "$P_a p$" (que corresponde a posibilidad) significa que "es posible, hasta donde sabe a, que p". Para la creencia, "$B_a p$" significa "a cree p" y "$C_a p$" que "p es compatible con todo lo que cree a". En su análisis del conocimiento, Hintikka discute algunos criterios de consistencia, de los cuales el más importante es el siguiente:

A_K: Si un conjunto de oraciones L es consistente y si "$K_a p_1$", "$K_a p_2$"..."$K_a p_n$" y "$P_a q$" son todos miembros de L, entonces el conjunto ("$K_a p_1$", "$K_a p_2$"..."$K_a p_n$", q) es también consistente.

En otras palabras, "si es consistente que yo diga que es posible, según mi conocimiento, que q es el caso, entonces tiene que ser posible que q sea el caso sin invalidar ninguno de mis requerimientos al conocimiento".[55] La palabra importante aquí es *requerimiento* al conocimiento. Si sólo requiriéramos que q fuera el caso sin invalidar lo que conozco, entonces el criterio siguiente, más débil, sería correcto:

B_K: Si un conjunto de oraciones L es consistente y si "$K_a p_1$", "$K_a p_2$"..."$K_a p_n$" y "$P_a q$" son todos miembros de L, entonces el conjunto ($p_1, p_2 ... p_n, q$) es también consistente.

La diferencia entre A_K y B_K es que la primera es un criterio de conocimiento como creencia verdadera *justificada* y la segunda un criterio de conocimiento como creencia verdadera, solamente. Digamos que si resultara que q, si es verdadera, minara mi requerimiento para conocer, digamos, p_1, por ejemplo, demostrando la falta de confiabilidad de la evidencia sobre la cual baso mi requerimiento de saber que p, entonces el sistema cognitivo comprometido en los antecedentes de los criterios sería inconsistente si el conocimiento se toma como creencia verdadera justificada, pero no si se toma como creencia verdadera. Hintikka propone como un criterio más, y muy obvio, el siguiente:

C_K: Si el conjunto de oraciones L es consistente y si "$K_a p$" es un miembro de L, entonces el conjunto $L+(p)$ es también consistente.

A partir de A_K y C_K obtenemos de inmediato:

D_K: Si el conjunto de oraciones L es consistente y si "$K_a p_1$","$K_a p_2$"..."$K_a p_n$" y "$P_a q$" son todos miembros de L, entonces el conjunto $(p_1 p_2 ... p_n$,"$K_a p_1$","$K_a p_2$"..."$K_a p_n$",$q)$ es también consistente.

Cuando vamos del conocimiento a la creencia, obtenemos criterios A_B, B_B, C_B y D_B, reemplazando en todas partes "$K_a p$" por "$B_a p$" y "$P_a q$" por "$C_a q$". De estos criterios, C_B claramente no es válido, pues en verdad queremos permitir creencias falsas. Los criterios A_B y B_B son válidos si D_B lo es, de manera que podamos concentrarnos en el argumento de Hintikka acerca de la necesidad de imponer el criterio de consistencia D_B, que puede ser expresado verbalmente: "Si algo es compatible con todo lo que usted cree, entonces debe ser posible que este algo resulte ser el caso juntamente con todo lo que usted cree sin hacer necesario que abandone ninguna de sus creencias."[56] Hintikka aporta un ejemplo asombroso e ingenioso para demostrar la necesidad y las implicaciones de este criterio. Véanse las siguientes oraciones:

Los laputanos creen que van a ser atacados por los ruritanos.

Los laputanos creen que los ruritanos los atacarán sólo si el ataque es por sorpresa,

en donde la segunda oración debe leerse

Los laputanos creen que los ruritanos los atacarán sólo si no creen que van a ser atacados.

Esto significa que tenemos un par de oraciones de la forma $(B_a p, B_a (p \supset \overline{B_a p}))$. De acuerdo con Hintikka este sistema de creencias es tal que debe ser intuitivamente caracterizado como inconsistente y es difícil no estar de acuerdo. Para un criterio que haga que este par resulte ser inconsistente, necesitamos D_B; ambas A_B y B_B son muy débiles. En el criterio D_B es sencillo ver que el conjunto $(B_a \overline{p}, B_a (\overline{B_a p} \supset p))$ también resulta ser inconsistente, mientras que los siguientes sistemas son ambos consistentes: $(B_a p, B_a (\overline{B_a p} \supset p))$ y $(B_a \overline{p}, B_a (p \supset \overline{B_a p}))$. Todo esto está bien en línea con nuestra intuición. Una implicancia interesante es la que se sigue: que una sola creencia $B_a (\overline{B_a p} \equiv p)$ es inconsistente en el criterio D_B. Porque si $B_a (\overline{B_a p} \equiv p)$, entonces una de las siguientes tiene que valer:

$(B_a (\overline{B_a p} \equiv p), B_a p)$ o $(B_a (\overline{B_a p} \equiv p), \overline{B_a p})$.

El primero de estos conjuntos es inconsistente porque el conjunto $(B_a (p \supset \overline{B_a p}), B_a p)$ es inconsistente, como se acaba de demostrar. El segun-

do conjunto puede ser reescrito como $(B_a(\overline{B_a p} \equiv p), C_a \bar{p})$. Si éste es consistente, entonces de acuerdo con D_B el siguiente tiene que ser también: $(B_a(\overline{B_a p} \equiv p), \overline{B_a p} \equiv p, \bar{p})$, pero del último deducimos de inmediato $(B_a(\overline{B_a p} \equiv p), B_a p)$, que nuevamente es el primer conjunto inconsistente. Los laputanos no pueden creer consistentemente que van a ser atacados si y sólo si no creen que van a ser atacados. Ellos pueden, sin embargo, contemplar consistentemente el siguiente sistema de creencias: $(B_a(B_a p \supset \bar{p}), B_a(B_a \bar{p} \supset p))$, ya que esto, mientras es incompatible con ambos $B_a p$ y con $B_a \bar{p}$, es compatible con $\overline{B_a p}$, es decir con $C_a \bar{p}$. Si creemos que el enemigo hará lo contrario de lo que creemos, si realmente creemos algo, entonces nos encontramos en terreno seguro sólo si no creemos nada.

Las últimas observaciones nos recuerdan el hecho de que la lógica de las creencias, tal como la de otras modalidades, tiene tres valores. Podemos creer que p, podemos creer que $\bar{p}$ o puede ser que no tengamos creencia alguna acerca de p. Podemos negar que $K_a p$ sin afirmar que $K_a \bar{p}$ no es difícil de aceptar, pero sugiero que muchas personas tienen una compulsión psicológica a creer que p o $\bar{p}$ una vez que la cuestión de p ha sido planteada. Una fuente importante de confusión puede ser que la lógica de la acción con frecuencia tiene dos valores, incluso si la lógica de las creencias sobre la cual se basa la acción tiene tres. Puedo creer que va a llover, puedo creer que el sol va a brillar o puedo no tener idea alguna acerca del clima, ni siquiera una de probabilidad subjetiva de distribución. De otro lado, la elección de llevar un paraguas o no está perfectamente clara, de manera que no hay una tercera opción. Es perfectamente consistente creer que, si llevo un paraguas, habrá sol, y que lloverá si lo dejo en casa; que la necesidad de un paraguas se dará si y sólo si no llevo un paraguas. Por otra parte, es inconsistente creer que lloverá si y sólo si no creo que llueva.

Una persona puede llevar su paraguas o dejarlo en casa, sin ninguna idea acerca del clima, actuando en cambio sobre la base de principios generales tales como razonamiento *maximin* o *maximax*, por ejemplo, actuando como si lo peor o lo mejor ocurriera seguramente. Puede también dejar o llevar el paraguas por alguna creencia sobre el clima. Estos dos casos pueden considerarse conjuntamente bajo el encabezamiento *creencia revelada* (en una analogía con la preferencia revelada), es decir, actuando como si creyera que p. Escribiendo "$RB_a p$" por "a actúa como si creyera que p", no hay inconsistencia en la creencia $B_a(\overline{RB_a p} \equiv p)$ ni, supongo, en la creencia revelada $RB_a(\overline{RB_a p} \equiv p)$, aunque no está claro para mí cómo actuar como si se creyera que la necesidad de un paraguas surgiría si y sólo si uno no lo lleva. Tampoco hay una inconsistencia en $(RB_a p, RB_a \bar{p})$. Alguien puede ser totalmente ignorante y no creyente en cuanto al clima, y sin embargo llevar el paraguas (actuando como si creyera que va a llover) y también bajar el toldo (actuando como si creyera que el sol saldrá en su ausencia). No hay inconsistencia cuando se toman precauciones contra dos hechos mutuamente excluyentes, incluso si uno no puede creer consistentemente que ambos puedan ocurrir. (En la siguiente sección volveré brevemente sobre el papel de la creencia revelada en la investigación científica.)

Un ejemplo más elaborado puede ser útil en este momento: el llamado *problema de Newcomb* enunciado por Robert Nozick.[57] Aquí se le dan las siguientes instrucciones a una persona:

> En la mesa, frente a usted hay dos cajas A y B. La caja B contiene 100 dólares. La caja A contiene, o bien,0 dólares o 1.000 dólares. Puede elegir abrir A y tomar como premio lo que está dentro o abrir A y B y tomar como premio lo que se encuentra dentro de las dos. El dinero fue puesto dentro de A antes que usted elija, por un Ser que en el pasado ha acertado un millón de veces ininterrumpido qué elección harán otros sujetos (dadas estas mismas instrucciones) luego de haber ubicado el dinero, y él ha colocado el dinero consistentemente de manera que todos los que han abierto la caja A han obtenido 1.000 y todos los que han abierto ambas cajas han obtenido 100 dólares solamente. En otras palabras, cuando el Ser prevé que alguien va a abrir ambas cajas, no pone nada en la caja A; si no, pone 1.000 dólares en la caja A y hasta ahora siempre ha tenido éxito en sus anticipaciones. Ahora elija usted.

En esta forma de enunciar las instrucciones (y en cualquier otra forma que imagine) hay una inconsistencia obvia, en que al millón de personas que lo precedieran en esta situación no se les puede haber dicho con verdad que el Ser ha hecho un millón de adivinanzas correctas seguidas. No hay ningún problema en la noción de un Ser que pueda predecir las elecciones de las personas a quienes se les ha hablado acerca sólo del dinero sin contarles la historia del Ser omnisciente: él predecirá (correctamente) que abrirán ambas cajas. Pero la idea del Ser prediciendo la elección de personas que creerán en dicho Ser es más problemática, porque ¿cómo llegaron estas personas a creer en la existencia de dicho Ser, en primer lugar? ¿Seguramente, no al darse cuenta de sus exitosas predicciones en el caso de las personas que no creían en su existencia? Permítasenos ignorar esta complicación y suponer solamente que estamos tratando con un individuo que cree en la existencia de dicho Ser, aun cuando sus bases para ello sean menos que adecuadas.

El dilema de tal individuo es obvio. De un lado él sabe que el dinero ya está en la caja A, así que no pierde nada si abre ambas cajas. De otro lado las predicciones exitosas del Ser indican que de alguna manera tiene posibilidad de prever la elección. Como el individuo cree en la historia del Ser, estaría dispuesto a apostar dinero por cualquiera que estuviera en la misma posición, obteniendo sólo 100 dólares si abriera las dos cajas, así que ¿por qué no aplicar el mismo razonamiento para sí mismo? Ahora, sea "p" la afirmación "Hay 1.000 dólares en la caja A" y "q" la afirmación "El individuo a abre ambas cajas". Permítasenos hacer una serie de suposiciones. Primero, suponemos que el individuo desea tener un sistema de creencias consistente. Segundo, suponemos que es racional, así que $B_a p \supset q$. Tercero, suponemos—no muy extravagantemente— que él mismo se cree racional, así que $B_a(B_a p \supset q)$. Finalmente hemos supuesto que él cree la historia del Ser omnisciente, así que $B_a(q \supset p)$. Las dos últimas creencias implican B_a

($B_a p \supset p$), que se mostró anteriormente como incompatible con $B_a p$. Dado, entonces, que la segunda y tercera suposiciones son bastante razonables, la persona no puede creer consistentemente que la historia del Ser es verdadera, y que hay dinero en la caja A. De otro lado, abrir la caja A solamente no tendría sentido a no ser que creyera que p. Se desprende que— bajo las presunciones establecidas— él tendría que abrir ambas cajas.

Para el lector que desea una confusión compleja más bien que una simple podemos hacer un breve comentario acerca del siguiente razonamiento de Anatol Rapoport: "Conociendo la gran exactitud del Ser para predecir mi elección, engendro una creencia de que muy probablemente voy a tomar la decisión correcta. Lo que está dentro de la caja no está bajo mi control, pero lo que yo creo que está allí sí lo está. Actuaré de manera que engendre una creencia de que tengo una oportunidad excelente de obtener los 1.000 dólares."[58] Ahora, en primer lugar (como se observó antes, rápidamente), no se puede *tomar la decisión de creer* algo para lograr un fin. Se cree algo porque se tienen razones para ello, no para alcanzar algo, aunque se puede fingir una creencia con algún objetivo.En segundo lugar, esto *a fortiori* es también válido para el doblemente imposible hecho de decidir tener un sistema inconsistente de creencias. La única salida posible, quizá, sería seguir a Pascal y establecer una serie de acciones cuyo resultado final (predecible) es mi creencia en lo imposible. En tercer lugar, sin embargo, Pascal no dijo que la creencia inducida por actuar como si creyera es la religión del "Yo creo porque es absurdo", que es lo que Rapoport nos llevaría a aceptar. Y, en cuarto lugar, ¿no sería el deseo de transformarse en el tipo de persona capaz de tener creencias inconsistentes un deseo contradictorio?

Creencias contradictorias: Festinger

El estudio más importante de lo que ocurre cuando las personas tienen sistemas inconsistentes de creencias es sin duda el trabajo de Leon Festinger sobre la disonancia cognitiva. En la definición formal de la noción, Festinger afirma que "[los elementos cognitivos] x e y son disonantes si no x se sigue de y".[59] Esta definición, sin embargo, resulta estar menos estrechamente relacionada con la inconsistencia epistémica que lo sugerido por la formulación. Los "elementos cognitivos" no son (o no sólo son) realmente cognitivos en sentido estricto. De acuerdo con Festinger, incluso una acción puede ser un elemento cognitivo, como cuando se dice que hay disonancia cognitiva cuando una persona continúa fumando sabiendo que fumar es peligroso. Un deseo o una preferencia puede también funcionar como un elemento cognitivo, como cuando una persona va a un concierto y se da cuenta de que quería ir al cine. La justificación de Festinger para el término"cognitivo" en estos casos es que si la persona

hace algo presumiblemente también sabe y *a fortiori* cree que lo hace; y de manera similar que "*a* desea *p*" implica "*a* sabe y *a fortiori* cree que desea *p*". Esto, aunque correcto, no es suficiente para hacer que, en el sentido de Festinger, la disonancia cogitiva resulte como las creencias incon-sistentes de Hintikka. No se puede derivar ninguna contradicción formal del conjunto de afirmaciones ("*a* cree que él fuma", "*a* cree que no debería fumar"). Tendríamos que hablar más bien de disonancia práctica en esos casos, o de debilidad de la voluntad. Es también claro que la frase "se sigue de" en la definición no debe interpretarse como una implicación lógica, sino más bien como una implicación "psico-lógica". El conocimiento de haber escogido el auto A y no el B es disonante con el conocimiento (adquirido luego de haberlo comprado) de que el auto B es en los aspectos principales mejor que el auto A: aquí no hay creencias contradictorias ni disonancia práctica, sólo un sentimiento de arrepentimiento.

Sin embargo, los sistemas de creencias inconsistentes (en el sentido de Hintikka) son una especie de la familia vagamente definida de sistemas cognitivos disonantes, y por lo menos algunas de las nociones, problemas y teorías que han evolucionado tienen alguna relevancia para nuestra discusión. La tesis general de la escuela de Festinger es la siguiente: "La existencia de la disonancia da surgimiento a presiones para reducir la disonancia y evitar crecimientos en la disonancia.[60] Comentaré algunos aspectos de esta tesis, en cuanto se relaciona con los sistemas de creencias contradictorios. Como una primera observación, daré algunos detalles sobre lo que se dijo anteriormente en la discusión acerca de la crítica de Popper a la dialéctica: el descubrimiento de contradicciones lógicas en una teoría es ciertamente un gran impulso para un cambio de teoría. En lógica y matemáticas esto pasa continuamente; ser contradictoria es, en verdad, la principal (aunque no la única) forma en que una teoría matemática puede ser falsa. En las ciencias empíricas las hipótesis lógicamente inconsistentes parecerían ser la excepción más bien que la regla. La hipótesis ergódica —una idea muy atractiva desde Descartes a Maxwell[61]— resultó ser contradictoria, pero no conozco muchos otros ejemplos sustanciales. De acuerdo con muchos críticos, el psicoanálisis y el marxismo son inconsistentes porque implican su propia falsedad, pero no creo que ésta sea una objeción válida. (Sería válida si se pudiera imputar a estas teorías las ideas de que cualquier teoría que pueda ser causalmente explicada es falsa y que todas las teorías pueden explicarse así, pero no creo que ninguna de estas imputaciones puedan sustentarse.) Las teorías empíricas usualmente se demuestra que son falsas por razones empíricas.

Podemos observar aquí, para referirnos a esto luego, que en el trabajo histórico las creencias que se revelan inconsistentes pueden jugar un papel genuino. Supongamos que quiero defender la tesis de que la importancia del hecho *x* (por ejemplo, la invención del ferrocarril) para el

fenómeno *y* (por ejemplo, el crecimiento económico norteamericano) fue más bien pequeña. Para probar este hecho, presumo que *x* no ocurrió y luego calculo el valor hipotético de *y*. Para este cálculo puedo necesitar el valor de alguna otra variable *z*, el cual (i) es causalmente afectado por la ausencia de *x*, (ii) es causalmente pertinente para *y*. Permítasenos asumir que *z* es pertinente para *y*, de dos maneras diferentes: podemos suponer, por ejemplo, que cuanto mayor es *v* y mayor es *w*, mayor es *y*, y que cuanto mayor es *z*, mayor es *v* y *menor* es *w*. Finalmente, presumimos que somos incapaces de encontrar el valor hipotético exacto de *z*, pero que tenemos confianza en que debe encontrarse en el intervalo (z_1,z_2). En ese caso me comportaré como si $(z=z_1)$ cuando calculamos *w* y como si $z=z_2$ cuando calculamos *v*; sin embargo, si el valor de *y* no difiere en mucho de su valor real, la hipótesis se corrobora. En el trabajo científico, el Abogado del Diablo puede tener creencias inconsistentes.[62]

Si estamos trabajando con creencias genuinas más bien que con creencias reveladas, la idea de que la inconsistencia induce al cambio puede parecer muy obvia o trivial para merecer comentarios posteriores. Pero, sobre la amplia y vaga definición de disonancia como cualquier clase de incompatibilidad lógica o emocional o "carencia de ajuste", se ha argumentado que las personas pueden realmente luchar por la disonancia, porque encuentran desagradable la consistencia total. Anthony Kenny[63] cita a Bertrand Russell, quien dice que "estar sin algunas de las cosas que se desean es una parte indispensable de la felicidad" y añade el comentario de que "en la medida en que Russell quiere ser feliz, él tiene, en conformidad con su dictum, que desear estar sin algunas de las cosas que desea". No creo que Kenny esté en lo cierto cuando ve una inconsistencia allí. Dada mi no posesión de las cosas $a_1, a_2...a_n$ (permítasenos llamar a esta no posesión ø) puedo preferir a_i a ø para todas las *i* y sin embargo preferir ø a $(a_1, a_2...a_n)$ — es decir, a la posesión simultánea de todas estas cosas — o al menos preferir $(a_1, a_2...a_{i-1}a_{i+1}...a_n)$ a $(a_1, a_2...a_n)$ para todas las *i* o para algunas *i*. (Lo que está presente aquí es la falacia de composición, de la cual se verá más en el capítulo 5.) Sin embargo, creo que está en lo cierto cuando dice que la correspondiente actitud epistémica es contradictoria. Es definitivamente inconsistente según los criterios de Hintikka creer que algunas de las cosas que yo creo son falsas. Yo también pienso que Kenny tiene razón cuando argumenta que la modestia debería prevenirme de afirmarme infalible, que sería exactamente lo que afirmaría si creyera, digamos, que todas las afirmaciones de este libro son verdaderas. Pienso que son verdaderas individualmente, pero creo que colectivamente hay algunas afirmaciones falsas. Individualmente, también podría creer que cada una de mis acciones presentes son moralmente correctas y creo también que más tarde encontraré que algunas de ellas fueron (son) incorrectas. Sólo deseo saber cuáles son. Si esto es inconsistencia (y lo es), entonces permítasenos guardarnos de la consistencia.

Un problema similar surge en la discusión de Festinger acerca de las posibles estrategias que las personas pueden emplear para reducir la

disonancia. Una de estas estrategias —quizá la más discutible[64]— es la exposición selectiva a la información. En la primera versión de la teoría de la disonancia cognitiva Festinger afirmó que "si una persona es conducida, por una u otra razón, hacia la expectativa de que una fuente de información producirá cognición, lo cual aumentará la consonancia, se expondrá a la fuente de información. Si la expectativa es que la cognición adquirida a través de esta fuente aumentaría la disonancia, la evitará".[65] Ahora, hay una paradoja evidente y latente en esta formulación, porque si una persona tiene ya bases para creer que una fuente dada de información va a contener información disonante con su estado mental, entonces él ya *tiene* información disonante con su estado mental y, por lo tanto, no tendría que tener un incentivo forzoso para evitar esa fuente.La información acerca de la información disonante con *X es* información disonante con *X*. Como lo ha enfatizado repetidamente Kenneth Arrow, la información como bien tiene las características peculiares de que "su valor para el comprador no es conocido hasta que tenga la información, pero entonces la ha adquirido sin costo alguno".[66] En el presente caso el valor es negativo, pero el problema es similar. Decir "creo que *p* y creo también que en un libro A hay información incompatible con *p*, aunque no sé qué información es esa" es inconsistente de manera bastante similar a como la admisión de falibilidad era inconsistente. En ambos casos pienso que de manera general y no concreta tengo falsas creencias, aunque no puedo precisar la contradicción.

Veremos que en sus últimos trabajos Festinger reconoció que sus afirmaciones más tempranas eran equivocadas, pero primero debemos comentar acerca de una forma ingeniosa en la que fue abordada esta dificultad en uno de los estudios sobre el cual Festinger basó su primer libro. Si uno pudiera encontrar una fuente privilegiada de información en la que la persona comprometida pudiera tener expectativas de encontrar información que redujera la disonancia o información que la aumentara sin que la expectativa en sí misma redujera la disonancia o la aumentara el problema se resolvería. Dicha fuente está obviamente haciendo propaganda, donde lo único seguro es que "siempre promueve el producto que anuncia";[67] se puede también considerar la "información" diseminada por los partidos políticos. Por lo tanto, si quiero afirmarme en la compra del automóvil A más bien que del B, yo sé *a priori* que obtendré lo deseado si leo los anuncios de los autos A mientras evito los del B, sin que este conocimiento sea equivalente (en sus efectos) a la lectura real. Ahora, no niego la existencia de personas que alquilan a otras personas para que les digan lo que quieren escuchar, pero sí quisiera cuestionar su racionalidad. El lector ávido de propaganda de productos que ya posee es, como el ejemplo de Wittgenstein, un hombre que compra varios ejemplares de un mismo matutino para asegurarse de que lo que dice es verdad;[68] podríamos invocar una analogía que utilizamos anteriormente, imaginando a una nación que intenta obtener reconocimiento diplomático de su propia colonia. En todos estos casos, un actor intenta derivar placer genuino de lo que recibe con su mano izquierda mientras, vigorosamente, trata de olvidar que el verdadero dador fue su mano derecha. Para una persona racional,

tanto el reconocimiento como la información tienen que derivar de alguna fuente independiente si se desea que reduzcan la disonancia.

En sus últimos trabajos, Festinger reconoció que "una vez que a la persona se le ha dicho que existe información contraria a su decisión, ya ha introducido información adicional—en el sentido de que es imposible, para él, obviarla ya que sabe que existe."[69] Cuando una persona posee conocimiento tan general y no concreto, dos tendencias opuestas trabajarán simultáneamente. De un lado, "bien podría haber una tendencia que obviara descubrir los detalles concretos de la información disonante con la decisión", pero de otro lado, "sin conocer los detalles concretos, es muy difícil refutar, reinterpretar y reducir la disonancia".[70] En los experimentos informados por Festinger estaban presentes ambas tendencias en proporciones variables, haciendo que el resultado neto fuera más bien ambiguo. Aquí quisiera añadir algunos argumentos generales acerca de la importancia de la primera tendencia. El impacto de las afirmaciones y creencias parece, en muchos casos, reducirse cuando tienen que ser expresadas en términos más bien generales. Mi creencia de segundo orden de que algunas de mis creencias de primer orden son falsas no entra de lleno en conflicto con estas creencias de primer orden, porque pertenece a un nivel diferente de generalidad. Decir de alguien "si yo diera mi opinión real de él, estaría sujeto a las leyes de difamación de carácter" no es colocarse uno mismo como sujeto de las leyes de difamación de carácter. Las cintas de Watergate mostraron que "no deseo saber los detalles" era una expresión recurrente de las divagaciones del principal actor. Yo puedo creer que la tortura, la detención ilegal, el chantaje, etcétera, no ocurren en las instituciones gubernamentales, sabiendo, al mismo tiempo, que algunas de las cosas que yo deseo que se hagan (y que se hacen) sólo pueden ser cumplidas por esos medios. Mientras yo no conozca *cuál* de estos medios es empleado, la inconsistencia es más soportable que si yo tuviera conocimiento al nivel concreto.

Antes de ir a la discusión de las contradicciones sociales en el capítulo 5, vuelvo brevemente sobre los objetivos fundamentales del presente análisis. He intentado mostrar, primero, que uno puede dar un significado preciso al término "contradicción real" en un sentido que no es susceptible de las objeciones analíticas corrientes. Segundo, argumento que esta noción no sólo puede precisarse (después de todo, abundan las precisiones sin importancia) sino que es extremadamente fértil también. Esto es así especialmente por la articulación entre las contradicciones y el cambio. En verdad creo que las contradicciones de la mente representan uno de los dos principales impulsores del cambio psíquico endógeno, siendo el otro la perversión de la intencionalidad por la causalidad. La necesidad de enfocar el cambio *endógeno* debe ser clara, porque la aproximación al cambio a través de la influencia exógena es sólo otra manera de escurrir el bulto. Algunas veces se debe decir: "El bulto se detiene aquí."

Notas

1. Me refiero al término *Widerspruch*, que no debe confundirse con *Gegensatz* (oposición). Por lo que sé, por ejemplo, Marx nunca se refiere a un *Widerspruch* entre el trabajo y el capital; sin embargo, tanto Louis Althusser (1965, pág. 104) como Maurice Godelier (1966) usan el término "contradicción" para esta relación.

2. Las afirmaciones contra las cuales estoy argumentando en las últimas tres oraciones se encuentran en Merleau-Ponty (1945, pág. 31). Véase también Taylor (1971, pág. 47).

3. Pörn (1970, págs. 34 y siguientes) propone una distinción similar entre contradicciones intrapersonales e interpersonales. En su discusión del ejercicio racional o consistente del poder, Pörn establece tres criterios de consistencia, el primero de los cuales es interno al que mantiene el poder y los dos últimos se ocupan de la relación entre varias personas que mantienen el poder y entre varios sujetos al poder.

4. Körner (1976, capítulo 16) propone una distinción similar entre las contradicciones prácticas y teóricas. Como Hegel, y a diferencia del actual capítulo, Körner también trata de resolver las relaciones de interdependencia y prioridad de las dos formas de consistencia. Sus resultados no llevan fácilmente a este enfoque, pero el lector interesado ciertamente debería consultar lo que él tiene que decir sobre el problema.

5. Para una exposición véase especialmente Lee (1936), quien da los textos griegos completos con traducción y comentario.

6. Este"es decir" está abierto a la controversia. Así G. Vlastos (1967) argumenta que es un error categorial suponer que la flecha debe estar o en movimiento o en reposo en cada momento, tal como lo es asumir que un punto debe ser derecho o curvo.

7. Una vez más, éstas son cuestiones controvertidas. Véase por ejemplo la distinción entre reposo e inmovilidad en Koyré (1971, pág. 33).

8. Una cita puede ser necesaria en este punto: "Algo se mueve, no porque en un punto del tiempo está aquí y allí en otro, sino porque en uno y el mismo punto temporal está aquí y no aquí, y en este aquí está y no está. Debemos aceptarles a los viejos dialécticos las contradicciones que ellos prueban en movimiento; pero lo que se deduce no es que no hay movimiento sino más bien que el movimiento es la Contradicción existente misma". (Hegel, 1961, vol. II, pág. 67).

9. Véase por ejemplo Whitrow (1961, pág. 141): "Lógicamente, una magnitud no cero no puede ser generada combinando una cantidad *finita* de magnitudes cero: *ex nihilo nihil fit.* Si, por lo tanto, un sonido audible puede producirse con el efecto combinado de una cantidad *finita* de 'sonidos inaudibles', entonces un conjunto finito de 'nadas' produjo 'algo', y por lo tanto nos enfrentamos con una contradicción entre la razón y la experiencia. La analogía con el argumento de que el movimiento no puede ser compuesto por un conjunto de inmovilidades es obvia". El uso del término "finito" (en itálicas aquí) por parte de Whitrow muestra, sin embargo, que la analogía está muy lejos de estar completa, ya que hay una cantidad infinita de instantes en el vuelo de la flecha, y de acuerdo con un escritor "el cálculo elemental nos enseñó que el producto de cero por infinito puede ser igual a cualquier valor finito" (Capek, 1972, pág. 336). Ahora el cálculo elemental ciertamente no nos enseñó nada de esto, pero sigue siendo verdad que hay una diferencia esencial entre los casos finito e infinito.

10. Así J. Laplanche (1970, pág. 6) explica la aparición de afirmaciones contradictorias en los escritos de Freud por el hecho de que estaba tratando con un objeto contradictorio, la mente humana. Véase también la siguiente discusión de una fuente posible de disonancia cognitiva: "Tercero, el error puede estar en las estrellas y no en nosotros mismos. La noción de una sociedad libre de contradicciones inherentes está quedando fuera de moda. Las cogniciones que reflejan un mundo social inconsistente... deben ser internamente contradictorias" (McGuire, 1966). En el capítulo 5 demuestro que la última afirmación está equivocada: la conjunción de los sistemas de creencias que son individualmente consistentes puede dar un todo inconsistente; para un individuo puede ser posible guardar su torta y comérsela, pero esto no es así para todos los individuos simultáneamente.

11. Popper (1940).

12. Pen (1959, págs. 91 y siguientes).

13. Wittgenstein (1953, págs. 32 y siguientes).

14. Véase Hull (1965) para una discusión de definiciones disyuntivas en biología.

15. Como me lo señaló Dag Prawitz, esto no sería verdad en un sistema tal como el "cálculo mínimo" de Johansson (1936).

16. Esta definición de los deseos contradictorios puede ser comparada con la discusión en Körner (1976) de las diferentes clases de contradicciones prácticas. En su explicación (especialmente Körner, 1976, págs. 96 y siguientes) hay tres variedades: oposición, discordancia e incongruencia entre actitudes prácticas. La oposición es bastante similar a lo que aquí se define como un deseo contradictorio, mientras que la discordancia está relacionada con lo que se discute en el texto como una debilidad de la voluntad. La incongruencia se define como un conjunto de actitudes de primer nivel. A esta tricotomía Körner le agrega otra, verbigracia, una distinción entre la imposibilidad lógica, empírica y práctica. Esto parecería dar un conjunto muy rico de posiciones contradictorias, pero la falta de ejemplos no triviales en la discusión de Körner hace que sea difícil ver cuánto uno puede obtener de este confoque.

17. Para esta distinción véase también Sen (1970, pág. 61): "Un hombre que juzga que 'el consumo hoy en día debería ser maximizado y que el consumo dentro de un año también debería ser maximizado' no está envuelto en... un conflicto lógico, ya que los dos juicios entran en conflicto sólo bajo circunstancias específicas (aunque plausibles)... en contraste, el buen hombre que desea que todos tengan un ingreso mayor que el promedio nacional parece tener un problema analítico de cierta magnitud."

18. Para dudas sobre la utilidad de la distinción analítico-sintético, véase Quine (1953, capítulo II) y Quine (1960).

19. Para este método de "refutación inmanente" véase la Introducción a la *Fenomenología del espíritu*.

20. En realidad, la relación entre amo y esclavo forma la matriz de las siguientes formas también. *Estoicismo* es el esclavo como amo; *escepticismo* es el esclavo como esclavo; *la conciencia desdichada* contiene tanto al amo como al esclavo dentro de sí misma, en una relación que ha sido comparada a la del superyó y el yo.

21. Hegel se inspiró especialmente en el *roman de culture* francés como *La Vie de Marianne*, de Marivaux. La mejor comparación parecería ser la de *A la Recherche du Temps Perdu*, que como el trabajo de Hegel es circular en estructura e interesado por la *génesis del pensamiento genético*. El proceso que permite que la mente entienda su propia historia es idéntico a esa historia.

22. Genovese (1974, pág. 146).

23. *The Prohibition*. Obsérvese que Donne, como Hegel luego de él, toma el punto de vista de que el propósito del amo se anula a sí mismo, de tal modo que no hay necesidad de que el esclavo sea su "propio vengador". Es apropiado reproducir aquí la primera estrofa del poema:

Cuidado con amarme.
Por lo menos recuerda que te lo prohibí;
No porque yo vaya a compensar mi pródigo derroche
De aliento y sangre, con tus suspiros y lágrimas,
Siendo para ti entonces lo que tú eras para mí.
Pero, tan gran alegría nuestra vida inmediatamente consume,
Entonces, para que tu amor, por mi muerte, no se frustre,
Si me amas, cuidado con amarme.

Esta es una idea mucho más convencional: si me amas demasiado me moriré de alegría y te privaré del objeto de tu amor. En el comentario en el texto sobre Sartre, Stendhal y Proust digo que desde un punto de vista moderno el amor es mejor por la misma paradoja que Donne aplica al odio; por cierto los dos casi no pueden ser separados.

24. Le debo esta observación a Jens Arup Seip; véase Seip (1958, págs. 34-35), quien atribuye la cita de Rousseau a sus *Lettres de la Montagne*.

25. Véase Elster (1975a., págs 171 y siguientes) para una discusión de este problema

en los escritos de Leibniz.

26. Véase nota 19 del capítulo 3.

27. Hintikka (1967).

28. Canarella y Tomaske (1975).

29. Genovese (1974, pág. 96).

30. Véase por ejemplo Milgram (1974, págs. 132 y siguientes), quien propone que no deberíamos distinguir entre *personas* obedientes y autónomas, sino más bien entre un *estado* obediente y uno autónomo de una persona dada. La analogía no está completa, porque en lo que Milgram llama el "cambio agencial" del modo autónomo al obediente el factor desencadenante está fuera del individuo mismo, mientras que los estados de ánimo cíclicamente cambiantes del amo no dependen de cambios en el medio ambiente. No hay ninguna *contradicción* en el análisis de Milgram, porque el marco de referencia es causal y no intencional.

31. Sartre (1943, pág. 95).

32. Sartre (1943, pág. 434).

33. Sartre (1943, pág. 708).

34. Sartre (1943, pág. 668).

35. Freud (1968).

36. Véase también Hyppolite (1966) y Lacan (1966) para comentarios sobre el texto en el que Freud niega realidad a la negación en el inconsciente.

37. Véase especialmente Bateson (1956); también Watzlawick y otros (1967).

38. Emily Dickinson. *Complete Poems* Nº 1560. El poema en su conjunto es así:

Ser olvidada por ti
Sobrepasa la memoria
De otras mentes
El Corazón no puede olvidar
A menos que contemple
Lo que declina
Fui mirada entonces
Sacada del olvido
Una sola vez
Para ser recordada qué -
Digno de ser olvidado
Es mi renombre

Podría decirse, tal vez, que en este ámbito, los tres versos citados en el texto adquieren un significado distinto; pienso, sin embargo, que se podría decir con igual justicia que la lectura ofrecida aquí permite una nueva y tal vez irónica interpretación del encuadre. "Si me vas a olvidar, cuidado con olvidarme". Véase también Elster (1976d).

39. Véase especialmente su *Pensée* 233, comentado en Elster (1976d).

40. La paradoja de Condorcet supone que una comunidad de tres individuos elegirá entre las alternativas (*a, b, c*) por voto mayoritario. Si la primera persona prefiere *a* a *b* y *b* a *c*, la segunda *b* a *c* y *c* a *a*, y la tercera *c* a *a* y *a* a *b*, se ve inmediatamente que hay una mayoría a favor de *a* en contra de *b*, de *b* en contra de *c* y de *c* en contra de *a*. Una condición suficiente que impide que surja este ciclo es la siguiente: para cualquiera de las tres alternativas (*a, b, c*) hay una alternativa, digamos *a*, tal que todos los individuos relacionados acuerdan que no es la mejor, o todos acuerdan en que es la peor, o todos acuerdan en que no es la media. Para detalles véase Sen (1970) o Riker y Ordeshook (1973).

41. Debido a Raiffa (1968, pág. 78).

42. En realidad la noción de un mejoramiento a muerte establece un vínculo entre las contradicciones de la mente y las contradicciones de la sociedad discutidas en el capítulo 5. En el segundo, un punto crucial es que lo que puede ser bueno para cada persona tomada individualmente puede ser desastroso para todos tomados colectivamente; en el mejoramiento a muerte el punto es que lo que puede ser bueno en cada momento tomado individualmente

es a veces desastroso para todos los momentos tomados colectivamente. En otras palabras, existe algo como la falacia de composición detrás de estas contradicciones mentales.

43. Von Weiszäcker (1971a, pág. 356) utiliza el término "preferencia" para esta tendencia, pero esto es claramente inadecuado si se lo toma literalmente. Es muy importante aquí distinguir entre el lenguaje intencional de las preferencias y el lenguaje causal de las tendencias; véase también apéndice 2 del capítulo 5.

44. Para lo cual véase Strotz (1955-56), Ainslie (1975) y Elster (1976d.).

45. En realidad las cuestiones son más complejas. Sin preferencias temporales, y dadas ciertas condiciones técnicas, puede no haber plan óptimo; para cada plan que incluye una cantidad positiva de consumo (por encima de la subsistencia) se puede encontrar otro plan mejor que incluya menor consumo, pero el plan que no incluya consumo es inferior a todos ellos. El punto es que el conjunto de todos los planes factibles es abierto y no contiene uno de sus puntos de acumulación, verbigracia, el plan de consumo cero. Véase Heal (1973) y Hammond y Mirrlees (1973) para una mayor discusión.

46. Véase Strotz (1955-56) para un análisis que quiebra el camino de este problema.

47. Agustín, *Confesiones*, VIII, vii.

48. Desafortunadamente, el análisis de Strotz de la planificación consistente era matemáticamente incorrecto, como fue demostrado por Pollak (1968). Dentro del marco de Pollak un ejemplo simple es el siguiente. Asumimos que estamos tratando con un período de tres años y con una cantidad K de mercaderías de consumo que deben ser colocadas durante el período. Establecemos C_1, C_2 y C_3 para las cantidades consumidas en los tres años; también dejamos que U_1 (C_1, C_2, C_3) sea la función de utilidad al comienzo del primer año y U_2 (C_2, C_3) la función de utilidad al comienzo del segundo. Si las preferencias temporales son consistentes, entonces la distribución hecha al comienzo del primer año de acuerdo con U_1 será tal que la pareja (C_2, C_3) incluida en esta distribución también maximizará U_2 (C_2, C_3), sujeto a $C_2 + C_3 = K - C_1$. Si, sin embargo, las preferencias temporales no son consistentes, se presentan tres posibilidades. El individuo puede actuar ingenuamente y cambiar sus planes al comienzo del segundo año; puede atarse (por ejemplo a través de una renta vitalicia) a la distribución que sea mejor de acuerdo con U_1; finalmente —si la primera opción es indeseable (porque el individuo es demasiado sofisticado) y la segunda imposible (por razones institucionales)— puede elegir la estrategia de la planificación consistente, sobre la base del siguiente argumento. Sabe que cuando ha elegido C_1 (no importa cómo), tendrá en su poder $K - C_1$ al comienzo del segundo año. También sabe que esta cantidad será distribuida a lo largo de los últimos dos años de acuerdo con U_2. Dados U_2 y $K - C_1$ puede determinar C_2 y C_3; establecemos que $C_2 = g(K - C_1)$ y $C_3 = K - C_1 - g(K - C_1)$. El sabe, es decir, que siendo la clase de persona que es, sólo secuencias de consumo de la forma $(C_1, g(K - C_1), K - C_1 - g(K - C_1))$ pueden simultáneamente satisfacer la restricción presupuestaria y su visión sofisticada de sus propias preferencias inconsistentes. Entre todas las secuencias posibles de esta forma entonces elige la que es mejor de acuerdo con sus preferencias al comienzo del primer año; en otras palabras: elige C_1 para maximizar U_1 $(C_1, g(K - C_1), K - C_1 - g(K - C_1))$. Esta estrategia satisface ambas, razón y pasión; razón porque la elección es sofisticada más que ingenua, pasión porque las preferencias instantáneas determinan las elecciones. Para algunas dudas sobre esta ingeniosa estrategia, véase Elster (1976d.).

49. Davidson (1969); véanse también los trabajos reunidos en Mortimore (1971).

50. Véase por ejemplo Davidson (1963; 1970).

51. Davidson (1969, pág. 111).

52. MacIntyre (1962).

53. Véase Singer (1963, págs. 58-59) para una crítica similar de B. Mayo.

54. Véase Husserl (1913, párrafo 25) para una exposición y crítica de esta interpretación.

55. Hintikka (1961, pág. 17).

56. Hintikka (1961, pág. 24).

57. Nozick (1969).

58. Rapoport (1975, pág. 619), algo modificado.

59. Festinger (1957, pág. 13).

60. Festinger (1957, pág. 31).

61. De acuerdo con Descartes "la materia debe tomar sucesivamente todas las formas de las cuales es capaz" (*Oeuvres*, comps. Adam y Tannery, vol. IX-2, pág. 216 = *Principes de la Philosophie*, vol. III pág. 47), donde "capacidad" probablemente se refiere al requerimiento de una "fuerza" constante (erróneamente pensada por Descartes como un requerimiento de una constante *mv* interpretada a escala). La afirmación correspondiente de Maxwell es "que el sistema, si se lo deja librado a sí mismo en su actual estado de movimiento, tarde o temprano pasará a través de cada fase que es consistente con la ecuación de energía": (citado de Farquhar (1964, pág. 75). El problema subyacente que da surgimiento a la contradicción aquí es que una curva continua que no se intersecta a sí misma no puede llenar un espacio multidimensional.

62. Por supuesto si el valor de *y* resultara ser muy diferente, uno podría tratar un enfoque más sofisticado, buscando el valor de *z* en el intervalo que maximiza *y* cuando se los inserta a los dos en el cálculo de *v* y en el de *w*. Esto, sin embargo, puede ser impracticable e insumir mucho tiempo comparado con el método analizado en el texto.

63. Kenny (1965-66). Para afirmaciones sobre la necesidad de la disonancia, véase también Singer (1966, pág. 55) y Leibniz (1875-90, vol. V, pág. 175): *L'inquiétude est essentielle à la félicité des créatures.* Véase también nota 19 al capítulo 3.

64. Véase Freedman y Sears (1965).

65. Festinger (1957, pág. 128), citado de una versión ligeramente aumentada en Festinger (1964, pág. 96).

66. Arrow (1971a, pág. 152).

67. Festinger (1957, pág. 49).

68. Wittgenstein (1953, pág. 94).

69. Festinger (1964, pág. 82).

70. Festinger (1964, pág. 82).

5

Contradicciones de la sociedad

La progresión de la *Fenomenología del espíritu* va del modo de contradicción intrapersonal al interpersonal. De manera similar, el itinerario intelectual de Jean-Paul Sartre va del análisis de las contradicciones psicológicas en *El ser y la nada* al de las contradicciones sociales en *Crítica de la razón dialéctica*. Para Hegel, la transición de un tipo de contradicción a la siguiente era una parte esencial de su filosofía. En el caso de Sartre, la articulación es mucho menos evidente. Muchos se han preguntado si los últimos análisis de Sartre, en realidad, suplantan o complementan a los primeros. Mi punto de vista es que los dos son compatibles, pero que ninguno implica o presupone al otro. Es una relación parecida a la que hay entre este capítulo y el anterior. Pienso que sería posible, finalmente, elaborar una teoría que ligara las contradicciones psicológicas a las sociales en un marco consistente, así como debería ser posible ofrecer una exposición integrada de los deseos y las creencias contradictorios. No me he sentido apto para hacerlo, por lo cual los análisis presentes deben verse como ladrillos para esa construcción y pasos hacia una síntesis ulterior. En el apéndice 2 de este capítulo intento un avance en esa dirección.

La estrategia del presente capítulo es la siguiente. En una primera sección discuto más o menos extensamente la falacia de composición, cuya importancia para las ciencias sociales ha sido con frecuencia casualmente reconocida, pero nunca —según creo— colocada como objeto de un análisis sistemático. En la segunda sección defino una diversidad de contradicciones sociales, a la que —siguiendo a Sartre— llamo *contrafinalidad*: las consecuencias no intencionales de acciones no coordinadas. Ligada a la contrafinalidad hay otra variante fundamental, que conforma la tercera sección, a la que llamo *suboptimalidad*, tal como apareció en el Dilema del Prisionero. En la cuarta y última sección intento unir estas nociones en una *teoría del cambio social*, argumentando que las dos variedades de contradicciones y el cambio político-económico están —aunque de modo complejo— íntimamente articulados. Adiciono dos apéndices: en el primero bosquejo —en forma sumamente resumida— la teoría de los

juegos, para el lector no familiarizado con el pensamiento estratégico; en el segundo bosquejo algunas implicaciones de los dos capítulos sobre las contradicciones en la teoría del comportamiento racional.

Para dar una noción preliminar, estableceré al inicio un ejemplo transparente y sencillo, tomado de la economía keynesiana. Tal como han dicho los discípulos radicales de Keynes una "paradoja esencial del capitalismo" es que cada capitalista desea bajos salarios para sus trabajadores (esto le significa mayores ganancias) y altos salarios para los trabajadores de todos los demás capitalistas (lo que hace que haya una gran demanda).[1] *Es* posible para todo capitalista, tomado individualmente, tener su torta y comérsela, pero no todos los capitalistas simultáneamente pueden estar en esta agradable posición. Sería cometer la falacia de composición concluir, del caso individual, el colectivo. La paradoja del capitalismo está, también, en la teoría keynesiana de las crisis y la contrafinalidad. Dada una reducción de la demanda, exógenamente inducida, cada capitalista intenta reaccionar frente a la caída, reduciendo los salarios de sus trabajadores. Desde el punto de vista del actor individual, es una medida racional, pero, cuando todos los capitalistas lo hacen colectivamente, el resultado es una reducción mayor del poder adquisitivo de los salarios y, por lo tanto, de la ganancia, con lo que se obtiene un resultado contrario al pretendido. Los capitalistas pueden, finalmente, entender esto, (y comprender que todos lo entienden, etc.) en cuyo caso la situación se convierte en un juego, cuya solución, sin embargo, conduce al mismo comportamiento anterior: cada capitalista reducirá los salarios, por temor a ser el único que no lo hace, mientras que antes lo hizo con la esperanza de ser el único que lo hacía. La suboptimalidad de la situación produce una paradoja y tensión; es duro aceptar la miseria, cuando todo el mundo podría estar mejor si todo el mundo se comportara diferente. El resultado, dadas ciertas condiciones estructurales, es una presión hacia la organización, la acción colectiva o la intervención gubernamental: un caso de *contradicciones que inducen al cambio.*

Antes de entrar en los detalles de estas ideas, permítaseme señalar cuál es, en mi opinión, su importancia para el marxismo. Argumentaré que las más importantes apariciones del término "contradicción" en los trabajos de Marx se ajustan bastante bien a mi análisis. En cuanto a los criterios para establecer lo que es importante, sugiero que sobre base del sentido común se acepte (a) que una aparición es más importante que otra si se encuentra más frecuentemente en el corpus y (b) que una aparición en el título de un capítulo o de una subsección de un capítulo es más importante que si aparece dentro del texto, ambos criterios sujetos a la cláusula *ceteris paribus*. De acuerdo con esto, se destacan tres casos: "la contradicción en la fórmula general del capital", discutida en el capítulo 5 (de la edición inglesa) del primer volumen de *El capital*; la "contradicción de la ley de la tasa decreciente de la ganancia", tema del capítulo 15 del tercer volumen; la "contradicción entre las fuerzas de producción y las relaciones de producción", a las que se refiere en muchos textos, especialmente *La ideología*

alemana y los escritos del período 1857-61. Esto será examinado, respectivamente, en la primera, segunda y tercera secciones siguientes.

La falacia de composición

La literatura filosófica y sociológica sobre la inferencia válida contiene muchas referencias a esta falacia. Daré primero una visión general sobre diversas inferencias falaces que parecen sustantivamente relacionadas entre sí, algunas de las cuales han sido llamadas "la falacia de composición". Escogeré, entonces, una de estas falacias y propondré que se la llame *la* falacia de composición. Hay, desafortunadamente, diversas razones que van en contra de esta propuesta. En primer lugar, hasta donde conozco, nadie se ha referido a esta falacia como de composición. En segundo lugar, a *la* falacia en cuestión no corresponde ninguna falacia de división, tradicionalmente antitética a la de composición. La antítesis de la falacia de composición es, de hecho, una inferencia válida y no hay ningún otro candidato obvio. Sin embargo, pienso que la estructura de la falacia corresponde mejor al término "falacia de composición" que a cualquier otra; también, la importancia intrínseca de la falacia hace que sea útil darle un nombre; finalmente, la falta de acuerdo sobre la definición apropiada debilita al primer contraargumento mencionado. Estos son temas controvertidos y no estoy llamando a pelear acerca de los términos. Si al lector no le gusta mi terminología, espero que centre más bien su atención sobre la sustancia de lo que digo. He aquí algunos ejemplos de falacias que, en cierto sentido, se puede decir que involucran composición.

(1) La interpretación más común en la literatura filosófica es la deducción a partir de"Cada miembro de este grupo es *F*" de "Este grupo de individuos es *F*"; la falacia de división es, entonces, tomada como la deducción antitética. En muchos casos, la falacia gira sobre la ambigüedad del nombre y la propiedad de "*F*"."Fuerte" no significa lo mismo cuando se aplica a un ejército que cuando se aplica a los individuos que forman el ejército. Por lo tanto, inferir hacia arriba o hacia abajo un tipo de fuerza de la otra constituirían instancias de las falacias de composición y división. En otros casos, la falacia surge, incluso, cuando no se encuentra ninguna ambigüedad. La palabra "indeciso" presumiblemente significa lo mismo cuando se aplica a los individuos que a los grupos: "incapaz de llegar a una decisión". Sin embargo, muy probablemente, un jurado indeciso no está compuesto por miembros indecisos, más bien lo contrario.[2] Es posible dar una caracterización independiente de las propiedades que permiten una o ambas inferencias, pero no me detendré sobre ello aquí.

(2) En vez de la oposición miembro-conjunto, la falacia de composición es a veces definida a través de la distinción parte-todo, como cuando inferimos que la silla es liviana del hecho de que una de sus partes es liviana.

Nuevamente aquí la falacia se debe a una ambigüedad en la palabra "liviana". Este término, de acuerdo con análisis recientes, debería definirse como "más liviano que la mayor parte de los objetos de su clase". Como cada objeto concebible pertenece a una infinidad de clases, hay una ambigüedad inherente que determina un campo muy abierto a las inferencias falaces. Si, por ejemplo, cada parte de una silla está situada en la clase de objetos de un mismo volumen, mientras que cada silla está situada en la clase de las sillas, todas las partes pueden ser livianas y no el todo, y viceversa. No he podido encontrar ningún ejemplo en el que la falacia parte-todo descanse sobre términos no ambiguos y sospecho que en este caso la ambigüedad es una condición necesaria para que la inferencia no sea válida. Esta conjetura, si es verdadera, podría relacionarse con el hecho de que un miembro y el conjunto del cual forma parte pertenecen a diferentes niveles, mientras que la parte y el todo pertenecen al mismo nivel.

(3) La falacia ecológica y su conversa sin nombre podrían también ser candidatos a los títulos de falacias de división y composición, respectivamente. La falacia ecológica es la inferencia de una correlación individual a partir de una correlación en el nivel de la comunidad, como cuando una alta proporción de negros en una comunidad con una alta tasa de delincuencia nos lleva a concluir que los negros cometen más delitos que los blancos. Esto no es válido porque la correlación en el nivel de la comunidad puede deberse también a que los negros son las *víctimas* más frecuentes de los crímenes.

Un ejemplo aun más dramático[3] es el siguiente: del hecho de que la delincuencia juvenil y la demencia senil están correlacionadas al nivel de la comunidad no podemos concluir que se encuentren en los mismos individuos. La falacia conversa no se discute a menudo, tal vez porque los datos sobre las correlaciones individuales son difíciles de encontrar y deben inferirse de otros datos, más bien que servir de base para la inferencia; cuando existen datos sobre las correlaciones individuales, usualmente permiten una inspección directa de la comunidad. Como ejemplo de la conversa de la falacia ecológica podemos revertir el problema negros-crimen. Si supiéramos que a escala nacional los negros cometen más crímenes que los blancos, no tendríamos que concluir que las comunidades con mayor cantidad de negros tienen mayor tasa de criminalidad, porque los negros en esas comunidades pueden sentirse más seguros y, por lo tanto, cometer menos delitos, o los blancos de esas mismas comunidades pueden sentirse más temerosos y, por lo tanto, cometer menos delitos. Una vez más, los datos subyacentes de la generalización nacional pueden haberse obtenido a nivel de la comunidad, de modo que la correlación a nivel de la comunidad puede ser inspeccionada directamente sin utilizarse la correlación individual a nivel nacional.

(4) En el capítulo 1 afirmé que la falacia de composición es la inferencia "$(\forall x)(M(Fx)) \supset M(\forall x)(Fx)$": lo que es posible para un solo indivi-

duo tiene que ser posible para todos simultáneamente. Esta, entonces, es la inferencia que propongo llamar falacia de composición. No se ha discutido mucho en la literatura,[4] pero argumentaré que es de fundamental importancia para la comprensión en las ciencias sociales. Para que esta inferencia sea una falacia debe ser posible mostrar casos donde el antecedente es verdadero y el consecuente, falso. Decir que el consecuente es falso implica decir que "$(\forall x)(Fx)$" es imposible. En ese caso, ¿en qué sentido puede ser contradictoria la atribución de una propiedad a todos los miembros de una clase? Excluimos de inmediato el caso (discutido en el capítulo 4) donde "Fx" es contradictorio en sí mismo, puesto que ello falsaría también al antecedente. La propiedad F tiene que ser en algún sentido *no universalizable* por razones que, aproximativamente, pueden ser clasificadas como lógicas, conceptuales y causales.

Es muy importante, en muchos contextos, la noción de propiedades no universalizables. Es también muy ambigua. Además de la universalizabilidad afirmada en la falacia de composición, tenemos la universalizabilidad afirmada en la siguiente inferencia: $(\exists x)(M(Fx)) \supset (\forall x)(M(Fx))$: lo que es posible para algún individuo es posible para cualquier individuo. En la teoría ética, en la que la necesidad y la posibilidad son interpretadas como obligación y permisión, esta inferencia es crucial: si algo se le permite a alguien, debe permitírselo a todos. Quizá más importante es el siguiente tipo de universalizabilidad ética: lo que es obligatorio para alguien es obligatorio para todos: $(\exists x)N(Fx)) \supset (\forall x)N(Fx))$. Edna Ullman-Margalit ha mostrado recientemente[5] que tanto la inferencia como la falacia de composición entran de manera esencial en el denominado *argumento de generalización*[6] de la teoría ética. Sea "U" el operador "no es deseable que", "O" el operador deóntico "es obligatorio que" y que "Fx" represente "x hace F". El argumento de generalización es, entonces, deducido de la siguiente manera:

GC: $U(\forall x)(Fx) \supset (\exists x)(O(\overline{Fx}))$
GP: $(\exists x)(O(Fx)) \supset (\forall x)(O(\overline{Fx}))$

GA: $U(\forall x)(Fx) \supset (\forall x)(O(\overline{Fx}))$

La premisa GC establece que si las consecuencias de que todos hagan F pudieran ser no deseables, entonces alguien tendría que hacer no F (o no hacer F). La premisa GP es el argumento de universalizabilidad recién presentado: si es obligatorio para alguien hacer no F, entonces es obligatorio para todos hacer no F. La conclusión es el argumento de generalización (muy estrechamente relacionado al imperativo categórico de la filosofía de Kant): nadie debería cometer una acción que pudiera tener consecuencias desastrosas si la cometiera todo el mundo. Ullman-Margalit comenta, muy convincentemente, que aunque la deducción es válida, la

premisa GC es inaceptable. Más bien debería reemplazarse por

AGC': $U(\forall x)(Fx) \supset O(\exists x)(\overline{Fx})$

La diferencia entre las consecuencias de GC y AGC' es la diferencia entre la modalidad *de re* y la modalidad *de dicto*. GC afirma que hay alguien (digamos Jones) que debería abstenerse de *F*, mientras que AGC' sólo dice que debería ser el caso que alguna persona (no especificada) se abstuviera de *F*. Si fuera desastroso que todos hicieran *F*, entonces, seguramente, debería ser el caso que no todos hicieran *F*, pero más que esto no podemos decir. Escoger un individuo y decir que *él* no debería hacer *F* sería arbitrario. Si alguien acepta AGC' y del consecuente de AGC' deduce el consecuente de GC obtendrá por supuesto GC, pero esta deducción es precisamente la falacia de composición en su forma alternativa y equivalente $N(\exists x)(Fx) \supset (\exists x)(N(Fx))$.

Volviendo a la noción de universalizabilidad afirmada en la falacia de composición tenemos que explorar el terreno en el cual se dice que una propiedad es no universalizable. Trataré fundamentalmente con las propiedades que son no universalizables por razones puramente lógicas. Ellas deben ser definidas racionalmente: "*Fa*" tiene que definirse por una expresión que envuelva conjuntamente *a* y las variables relacionadas *x,y,z...* y alguna relación. Los predicados simples, tales como tener pelo rojo o seis dedos, tienen que ser posibles de adscribir a todos simultáneamente si pueden adscribirse a cualquiera; pueden surgir contradicciones sólo si adscribimos a todos una propiedad que está definida en referencia a otras personas. La siguiente propiedad es un caso especial muy importante:

$$Fa \equiv (Ga \,\&\, (\forall x)((x \neq a) \supset \overline{Gx})) \equiv (\forall x)((x = a) \equiv Gx),$$

que es la propiedad de ser el único individuo que tiene la propiedad *G*. Todos los individuos pueden ser simultáneamente *G*, cualquier individuo podría ser *F*, pero todos no pueden ser simultáneamente *F*. Estableciendo "*Gx*" para "los trabajadores empleados por *x* tienen bajos salarios", reconocemos el ejemplo keynesiano afirmado al comienzo del capítulo.

Antes de dar más ejemplos de la falacia de composición en las ciencias sociales, ofreceré algunas observaciones sobre los casos en los que la propiedad en cuestión es no universalizable por razones causales o conceptuales . La distinción entre imposibilidad causal y lógica puede ser ilustrada por el siguiente ejemplo: es lógicamente imposible que todas las partículas de un gas tengan una velocidad por encima del promedio y causalmente imposible (pero lógicamente posible) que todas las partículas tengan exactamente la velocidad promedio, aun si puede presumirse que una partícula dada tenga una velocidad promedio sin entrar en conflicto con las leyes de la física. De hecho, Krönig dedujo la ley de Boyle a partir de la suposición de que todas las partículas tenían la velocidad promedio.[7] Si esta suposición fuera autocontradictoria, la deducción sería automática-

mente válida, como cualquier afirmación que puede deducirse de una contradicción. Sin embargo, éste no es el caso: la presunción es sólo una simplificación, de la cual puede deducirse un teorema que termina por valer en el sentido más general. La imposibilidad física está también presente en un ejemplo dado por Peter Geach:[8] de "Cada uno de estos libros puede ser leído en menos de media hora" no deberíamos concluir que "Es posible leer cada uno de estos libros en menos de media hora". La imposibilidad en este caso no se debe a ninguna relación intrínseca entre los libros, puesto que para comprender qué es leer un libro no tenemos que referirnos a otros libros. Más bien, la imposibilidad surge de la relación común de los libros con el contenedor temporal "media hora" y de ciertas limitaciones físicas de esa relación. Un problema similar está comprendido en la relación entre "Es posible traer transmisiones radiales locales a cualquier lugar de Noruega antes de 1980" y "Es posible traer transmisión radial local a todas las regiones de Noruega antes de 1980". Las restricciones físicas, tecnológicas y económicas hacen imposible concluir lo último de lo primero, a lo que podemos añadir que, por razones *políticas*, la falsedad de lo último hace surgir la falsedad de lo primero. Como observamos en el capítulo 3, en política ocurre a menudo el caso de que las propiedades no universalizables no pueden ser realizadas ni en los casos individuales. Un último ejemplo de la falacia causal de composición es la noción de hombre promedio.[9]

La versión de la imposibilidad conceptual de la falacia de composición es más interesante, pero también más difícil de analizar. Peter Winch[10] ha llamado la atención hacia una afirmación muy explícita de Hume acerca de la validez de la siguiente inferencia:

> En vano se pretende haber aprendido la naturaleza de los cuerpos por la experiencia anterior. Su naturaleza secreta y consecuentemente sus efectos e influencia pueden cambiar sin ninguna modificación de sus cualidades sensibles. Esto ocurre algunas veces, con respecto a algunos objetos: ¿Por qué no puede esto ocurrir siempre con respecto a todos los objetos? ¿Qué lógica, qué proceso de argumentación nos puede asegurar contra esta suposición?[11]

En contra de esto, podemos establecer una afirmación igualmente explícita de Wittgenstein:

> "Lo que algunas veces ocurre, puede ocurrir siempre".¿Qué clase de proposición es ésta? Es como la siguiente: si "*Fa*" es consistente "$(\forall x)(Fx)$" , es consistente."Si es posible que alguien haga un movimiento falso en algún juego, entonces sería posible que todos no hicieran ninguna otra cosa sino movimientos falsos en todos los juegos." Por lo tanto, estamos amenazados por la tentación de no comprender la lógica de nuestras expresiones, de dar cuenta incorrectamente de la utilización de nuestras palabras. A veces, las

órdenes no son obedecidas, pero ¿qué ocurriría si no lo fueran *nunca*? El concepto de "orden" perdería su objetivo.[12]

Ejemplos típicos de la imposibilidad afirmada por Wittgenstein (y por Winch) serían casos morales tales como todos rompiendo siempre promesas, casos epistemológicos tales como nuestros sentidos y memoria siempre engañándonos, y casos físicos tales como todos los objetos siempre rompiéndose y desintegrándose frente a nuestros ojos. Estos no son casos de imposibilidad lógica directa . Tampoco pienso que estos autores argumentarían que se deben a una imposibilidad causal de algún tipo. ¿Qué son, entonces? En mi opinión, sólo los casos morales son ejemplos genuinos de esa noción esquiva de imposibilidad conceptual. El problema de la ruptura de la promesa universal puede ser visto a la luz del imperativo categórico que afirma que uno no debería actuar siguiendo una máxima si uno "no puede desear" que llegara a convertirse en una ley general. Kant da varios ejemplos para explicar esta exigencia de universalizabilidad. De estos ejemplos, uno —negarse a ayudar a los otros— es no universalizable sólo en el débil sentido de que, si todos hicieran eso, las consecuencias serían desastrosas; la situación es, de hecho, un Dilema del Prisionero. (El lector recordará el argumento de Ullman-Margalit de que estas consecuencias desastrosas no dan un argumento para que todos ayuden siempre a los otros.) Estamos perfectamente capacitados para imaginar una situación de desconfianza, sospecha y egoísmo universales; tal mundo sería desagradable, pero no hay ninguna contradicción en el concepto.[13] Otro de los ejemplos de Kant es el no pago de los préstamos, el cual es idéntico en su estructura a la ruptura de las promesas. El no pago de los préstamos es no universalizable en el fuerte sentido de que no podemos siquiera imaginar —y *a fortiori* no desear racionalmente— una sociedad donde los préstamos no se pagaran nunca. En dicha sociedad, el concepto y la institución del préstamo estarían ausentes. Su pago está arraigado en la noción misma de que préstamo no es regalo. Esta idea puede extenderse a muchos otros casos de comportamiento significativo. Véanse por ejemplo las cartas que no llegan a destino. Puede concebirse que esto le ocurra a cualquier carta pero no es posible que les ocurra a todas, porque en ese caso las cartas no serían cartas sino pedazos de papel que, por razones misteriosas (quizá rituales), se arrojarían en buzones postales. Anticipando un punto que ahora veremos, no sólo los habitantes de dicho mundo carecerían del concepto de carta sino que *nosotros* tampoco diríamos que ellos enviaban cartas.

En cuanto a los ejemplos de física, asumo que Hume está en lo cierto. Sabemos lo que es que un objeto se desintegre frente a nuestros ojos y también sabemos lo que sería que todos los objetos se destruyeran frente a nuestros ojos. Winch argumenta que, en este caso, para nada deberíamos hablar de *objetos*, pero esta conclusión parece válida a lo sumo para los habitantes del mundo en cuestión. En un mundo en donde los objetos están permanentemente en formación y desintegración no tendríamos de dónde sujetar la noción de un objeto. En mi opinión, esto es tema más de la

psicología que del análisis conceptual o de la lógica. De todas maneras, los habitantes de un mundo (el nuestro, digamos) nada podría impedir a que hablaran de objetos que se desintegran en otro mundo (digamos, Marte). Si podemos hablar significativamente de algunos objetos que desaparecen en el mundo real, también podemos hablar significativamente de todos los objetos que desaparecen en uno de los mundos posibles (pero en *ningún* mundo sería posible que esas cartas nunca llegaran a su destino). Quizá podemos interpretar la imposibilidad presente en los objetos que constantemente se desintegran como una imposibilidad *causal*, comprendiendo una referencia no a las leyes de la física sino a las de la psicología. Podemos definir un mundo como accesible al nuestro (en el sentido técnico de la lógica modal) si nuestro mundo es concebible desde ese mundo.[14] Un mundo en donde constantemente se desintegran los objetos no podría (como un hecho psicológico) formar el concepto de un objeto; por lo tanto, nuestro mundo no podría concebirse desde ese mundo, el cual, *ipso facto*, no sería accesible desde el nuestro. Esto de nuevo implica que cualquier proposición que afirmara la posibilidad de que todos los objetos siempre se desintegraran sería falsa (en nuestro mundo). Un argumento similar es válido para los pretendidos casos de imposibilidad epistemológica.

Procedo ahora a dar un número de ejemplos de la falacia de composición en los casos en los cuales la propiedad pertinente es no universalizable, por razones puramente lógicas.

(a) Del hecho de que en cualquier sociedad dada pueda surgir una institución o técnica dada (según sabemos) por importación o préstamo, la teoría antropológica conocida como difusionismo concluye falazmente que esto puede explicar la innovación en todas las sociedades.

(b) Del hecho de que cualquiera puede depositar todo su dinero en el banco y obtener intereses se concluye a veces que todos pueden hacerlo simultáneamente.

(c) Del hecho de que cualquiera pueda retirar su depósito del banco cuando lo desee es tentador (pero falaz y desastroso) concluir que todos pueden hacerlo simultáneamente.

(d) Del hecho de que en una sociedad de propietarios de mercancías, un propietario individual pueda obtener ganancias comprando barato y vendiendo caro, algunos economistas concluyen que todos los propietarios de mercancías podrían obtener ganancias engañándose unos a otros. Esta es "la contradicción de la fórmula general del capital", a la cual se refiere Marx. Citando a Destutt de Tracy, quien dice que "los capitalistas industriales obtienen ganancias porque 'todos ellos venden por más de lo que les costó producir. ¿Y a quién venden? Unos a otros en primera instancia'", Marx comenta que "La clase capitalista, de conjunto, en cualquier país, no puede sobrepasarse a sí misma".[15]

(e) Del hecho de que la sobreproducción puede ocurrir en cualquier sector de la economía algunos viejos críticos del capitalismo (especialmen-

te Sismondi) concluyeron que la sobreproducción puede ocurrir en todos los sectores de la economía simultáneamente.[16] Obsérvese que ésta es una falacia sólo si la abstraemos de los argumentos keynesianos de la preferencia por la liquidez.

(f) Se pueden tomar diversos ejemplos de la lógica del análisis marginal, en la que siempre es tentador pensar que todo puede ser marginal. Un ejemplo conocido sería la forma en que se tiende a calcular el costo de los rubros a deducir de los impuestos, utilizando para todos los rubros el porcentaje de la tasa marginal del impuesto que realmente se aplica en un solo rubro. Un segundo ejemplo puede tomarse del análisis de James Coleman sobre el poder. El problema es cuánto poder atribuimos a un individuo que toma parte de un operativo conjunto controlado, como el voto colectivo. Como señala Coleman, la pregunta no debe contestarse utilizando la probabilidad del individuo de dar el voto definitivo, sobre la base de la presunción de que es igualmente posible que otros voten por sí o por no. En un modelo tripersonal, esto le daría a cada persona un poder relativo de 1/2 obteniendo un control total de 3/2, lo cual es contradictorio. La dificultad es que "Cada actor está asumiendo, de hecho, que él da el voto final. Pero esta presunción, evidentemente, no es válida, porque *cada uno de los tres no puede ser el último*".[17] Un tercer ejemplo fue discutido en el capítulo 4, donde vimos que Anthony Kenny argumentaba falazmente que es contradictorio querer estar sin algunas de las cosas que se desean. Un cuarto ejemplo es el siguiente:

> [Fábricas] que en los Estados Unidos producen componentes para ser ensamblados en otros países tienden a cobrar, al subsidiario que ensambla, el precio total mayorista por equipos "totalmente desmontados". En sí misma, esta práctica no ha sido generalmente considerada como explotadora. Pero cuando el artículo, como un farol o un paragolpes, es totalmente excluido del equipo antes del embarque, el precio del equipo ha sido reducido sólo por el costo marginal del artículo excluido. Algunas veces se ha observado, no siempre alegremente, que sobre esa base para fijar precios, el costo de cada contenedor embarcado, vacío de todo el contenido, generalmente excedería los 1.000 dólares.[18]

De la forma como está escrito este párrafo, parecería que el autor piensa que la última práctica mencionada es explotadora. La noción de *explotación del trabajo* será en verdad el último y más importante de mis ejemplos de este encabezamiento. En la teoría de la distribución del ingreso primero podemos diferenciar entre (la minoría de) los economistas que piensan que el trabajador debería de alguna manera "obtener lo que produce" (a cada uno de acuerdo con su trabajo) y aquellos que piensan que este principio es muy estrecho (¿qué pasa con los inválidos?). Segundo, hay diferencias dentro de la primera escuela, entre los que sostienen que el trabajador debería obtener el equivalente de su producción *promedio* y los que preferirían que se pagara de acuerdo con su producción *marginal*.[19] Quisiera sugerir que hay dos hechos institucionales propios del capitalis-

mo, que hacen que la interpretación marginalista de una justa distribución del ingreso parezca posible. En primer lugar, la existencia de un tipo de propietarios de capital parece justificar la noción de capital agregado como un "factor de producción" juntamente con el trabajo y con el mismo "derecho de remuneración"que el trabajo.[20] Utilizando el teorema de Euler para funciones productivas homogéneas de primer orden, sabemos que cuando ambos factores se pagan de acuerdo con su producción marginal, la remuneración total de factores agota el total del producto. En verdad, dada una función continua y diferenciable $f(K,L)$ de capital y trabajo, tal que $f(tK,tL) = t(f(K,L)$ para todas las t, el teorema de Euler afirma que $f(K,L) = f_K(K,L).K+f_L(K,L).L$, donde f_K y f_L son los derivados parciales de f con respecto al capital y el trabajo. Esta bella expresión de armonía de clases en lenguaje matemático parecería justificar la interpretación marginalista, puesto que si a ambos factores se les pagara de acuerdo con su *media* de producción el producto total se agotaría en más de dos veces. En segundo lugar, y ésta es la noción más pertinente para este punto de vista, las negociaciones salariales individuales hacen parecer irresistible a la teoría marginalista, porque cada trabajador puede ser tratado como si *él* fuera trabajador marginal. Con aparente justificación, el patrón puede decirle a cada trabajador: "Sin usted, la producción total se reduciría en una cantidad x; por lo tanto, lo más que puedo pagarle es el equivalente de x para poder mantenerlo empleado." Esto, aunque correcto para un solo trabajador, no puede ser verdadero para todos los trabajadores simultáneamente, porque entre ellos devoran todo el producto y no sólo el producto marginal multiplicado por el número de trabajadores.

Las dos razones aducidas para la aparente plausibilidad de la noción neoclásica de un ingreso justo como equivalente al producto marginal señalan los obstáculos que tiene que superar la clase obrera si reclama todo el producto como suyo. El paso de la negociación individual a la colectiva no es suficiente por sí solo, puesto que los obreros tienen que superar la alienación que hace que los bienes del capitalismo aparezcan realmente como propiedad de los capitalistas actuales más que como un producto de anteriores obreros. (En verdad pienso que esta utilización de los *principios resultado-fin* puede emplearse como definición de alienación.)[21] Esta doble solidaridad —entre todos los trabajadores en cada instante del tiempo y entre los trabajadores de sucesivas generaciones— es una condición necesaria y suficiente para que se disipe la ilusión de la teoría marginalista de distribución del ingreso. Me he detenido en este ejemplo porque me parece ejemplarizar en el sentido de mostrar la importancia *práctica* de la falacia de composición. *Divide et impera*; si cada hombre puede ser tratado como si fuera el hombre marginal, la clase que lo enfrente se reduce a polvo de individuos.

El lector puede haber observado que en el ejemplo (b) y (c) alguna cláusula como "siendo el único que..." tiene que presuponerse implícitamente, para que resulten falacias *lógicas* y no causales. Estos dos ejemplos

tienen también implicaciones obvias para la acción. Si todos depositaran simultáneamente su dinero en un banco, la economía se detendría inmediatamente y nadie obtendría interés alguno;[22] una corrida bancaria tiene también el efecto contrario de que motiva a los individuos a actuar. En verdad, muchas situaciones que pueden ser descritas a través de la falacia de composición dan también origen a las consecuencias no intencionadas que subsumo bajo el encabezamiento de contrafinalidad, que constituye el tema de la siguiente sección. Por definición (véase a continuación), la contrafinalidad comprende a la falacia de composición, pero muchas instancias de la falacia no tienen nada que ver con la contrafinalidad.

Contrafinalidad

Por este término entenderé las consecuencias *no intencionadas que surgen cuando cada individuo en un grupo actúa sobre una presunción acerca de sus relaciones con otros que, cuando se generaliza, produce la contradicción en el consecuente de la falacia de composición, siendo verdadero el antecedente de esa falacia.* Es claro que la contrafinalidad así definida es una especie de contradicción real, en el sentido de la tercera tesis sobre las contradicciones sostenida en el capítulo 4. El requerimiento de que el antecedente de la inferencia sea verdadero, es decir, que no haya nada contradictorio en la presunción individual tomada separadamente, garantiza la racionalidad individual. La irracionalidad colectiva surge sólo de la incompatibilidad de los sistemas de creencia. (Presumiblemente, el estudio de los sistemas de creencia incompatibles podría utilizar el mismo aparato de la lógica epistémica, como lo aplicamos en el capítulo anterior al análisis del sistema de una sola creencia contradictoria, pero no intentaré esto aquí..)[23] En los ejemplos siguientes, la presunción que da origen a la contradicción será siempre de la forma"siendo el único (el primero, el mejor, etc.) en hacer *x*". Esto, por consiguiente, es una abrevaitura de para "siendo entre los pocos (o el primero, el mejor, etc.) en hacer *x*". Sin embargo, antes de explicar estos ejemplos, creo que pueden ser útiles algunas observaciones introductorias sobre el concepto.[24]

La noción de contrafinalidad tiene muchos antecesores en la historia de las ideas —la mano invisible, la astucia de la razón, la alienación—. Entre quienes más han contribuido a la noción, me parece que debemos nombrar a Bossuet, Mandeville, Vico, Adam Smith, Hegel y Marx. Para discusiones recientes, pueden escogerse, quizás un poco arbitrariamente,los nombres de Oskar Morgenstern, Jean-Paul Sartre, Robert Nozick y Trygve Haavelmo.[25] En las filosofías de la historia de Bossuet y Vico,[26] la noción de consecuencias no intencionadas es ciertamente fundamental, pero hasta cierto punto —más en Bossuet que en Vico— la contrafinalidad es vista como un instrumento para la finalidad divina, de manera que su idea de causalidad social no define realmente un proceso inmanente. Si aceptamos la definición de Robert Nozick de las *explicaciones de manos invisibles* como una

presentación "explica lo que parece ser el producto del plan intencional de alguien como si no surgiera de las intenciones de nadie",[27] Bossuet y Vico ponen demasiado énfasis en la mano orientadora de Dios en la historia como para que se los pueda subsumir en esta categoría. Sin embargo, ambos trabajos contienen pasajes que, si se toman aisladamente, se leen, como Hegel o Marx, enfatizando la elección libre de los actores y, algunas veces, los resultados imprevistos de sus acciones.

Entre Bossuet y Vico está Bernard Mandeville, cuya importancia para la historia de las ciencias sociales es difícil de exagerar. Si se desea encontrar el origen de la noción de las funciones latentes, es necesario ir a Mandeville, no a Merton; acerca del concepto de la mano invisible, es necesario ir a Mandeville no a Adam Smith. Es también un pionero en su crítica a la falacia de composición, como puede verse en el elogio que le hace Keynes.[28] Su influencia ha sido en parte, subterránea, porque su lógica era muy rigurosa y sus conclusiones muy cínicas; probablemente era más leído que citado. Su lema "Vicios Privados, Beneficios Públicos" está mejor descrito en los siguientes pasajes:

> El miope vulgar casi nunca ve más allá de un enlace en la cadena de causas; pero aquellos que pueden ampliar su visión y darse el tiempo para contemplar la perspectiva de eventos concatenados, pueden, en cien lugares ver el *Bien* brotar y pulular a partir del *Mal*, tan naturalmente como lo hacen los pollos de los huevos.[29]

> Me jacto de haber demostrado que ni las cualidades afectivas ni los afectos amables son naturales al hombre, ni las virtudes reales que es capaz de adquirir por medio de la razón y la abnegación son los fundamentos de la sociedad, sino que lo que llamamos mal, tanto natural como moral, en este mundo es el gran principio que hace de nosotros criaturas sociables, la sólida base, la vida y soporte de todos los negocios y empleos sin excepción: que allí tenemos que buscar el verdadero origen de todas las artes y ciencias y que, en el momento en que el mal cese, la sociedad tendrá que dañarse, si no disolverse totalmente.[30]

Como lo he afirmado más detalladamente en otro lugar,[31] ésta es una aplicación sociológica de la teodicea de Leibniz. En verdad, Leibniz sostenía en el nivel metafísico que la existencia de algún mal es una *condición necesaria para* la maximización del bien, mientras su contemporáneo Malebranche sostuvo el punto de vista más tímido de que el mal es un *inevitable producto colateral del bien.* Sin embargo, en el nivel sociológico, Leibniz no asumió las consecuencias de su metafísica. El sostenía, por ejemplo, que el lujo debería ser aceptado como un lamentable subproducto de la vida próspera, mientras Mandeville fue a fondo y dijo que el lujo y la prodigalidad son las causas de la prosperidad.

Si se sustituye "interés personal" por "mal" en los pasajes arriba citados, tenemos a Adam Smith. El compartía con Mandeville la convicción de que la mano invisible no sólo es suficiente, sino necesaria para la

promoción del bien común:

> Al preferir apoyar la industria nacional a la extranjera [todo individuo] busca su propia seguridad; y si se orienta esa industria de manera tal que su producto sea del mayor valor, él pretende sólo su propia ganancia y está en ello, como en muchos otros casos, guiado por una mano invisible para promover una finalidad que no formó parte de su intención. Tampoco, el no ser parte de ello, es siempre lo peor para la sociedad. Si persigue su propio interés, frecuentemente promueve el de la sociedad más efectivamente que cuando realmente tiene intención de promoverlo. Nunca he conocido que aquellos que se dedicaron a los negocios por el bien público hayan hecho mucho bien.[32]

Se ha dicho a veces que Marx puso a Hegel cabeza abajo; más importante aún, por lo menos para nuestra finalidad, ambos pusieron a Adam Smith cabeza abajo. Para Hegel y Marx, la historia se desarrolla más bien a través de las consecuencias negativas de las acciones humanas, que de las positivas, a través de la contrafinalidad más que de la mano invisible.[33] Para Adam Smith, un caso paradigmático sería así: cada actor tiene una cantidad y_1 de y y hace x para obtener y_2; cuando todos hacen esto, todos obtienen y_3, en donde $y_3 > y_2 > y_1$. Hegel y Marx establecerían más bien $y_2 > y_3 > y_1$, o incluso $y_2 > y_1 > y_3$. Yo afirmo que siempre que vemos a la mano invisible ejecutar sus misterios, estamos tratando con una situación en la que (i) todos los actores tienen una estrategia dominante y (ii) lo que resulta del hecho de que todos los actores utilicen su estrategia dominante, domina todos los demás resultados. La primera característica implica que los actores realmente no toman en cuenta el comportamiento de los otros; puede ser, quizás, que mantengan creencias incompatibles unos acerca de otros, pero estas creencias no son esenciales para su elección. La segunda característica implica que el resultado es realmente favorable a todos. Como veremos enseguida, en contraste, la contrafinalidad puede darse cuando una de estas características está ausente. Sugeriría también que Nozick es inconsistente cuando se refiere a las teorías del ciclo económico o a la no provisión de bienes públicos como casos "explicados por la mano invisible". Ambos son casos de contrafinalidad (o quizá suboptimalidad, que discutiremos en la próxima sección) y nadie podría pensar que parecen "ser el producto de un plan intencionado de alguien". Nozick tiene razón en llamar la atención hacia el hecho de que los efectos benéficos de las acciones frecuentemente no son (y en verdad a veces no pueden serlo) los efectos que motivan estas acciones. Más adelante vuelvo sobre esta idea. Donde se equivoca es en los ejemplos que escoge para ilustrar su punto de vista.

Deberíamos hacer más precisa la noción de contrafinalidad antes de dar más ejemplos. Cuatro observaciones son pertinentes. En primer lugar, la noción de"actuar sobre la base de una presunción" puede ser interpreta-

da de manera psicológica, pero no necesariamente. Lo que estoy intentando es captar la noción de lo que es necesario para que la acción tenga sentido y esto no es siempre la misma cosa que el sistema de creencias explícito de los actores. Por ejemplo, en el modelo de la telaraña (que discutiremos más adelante) cada granjero actúa sobre la base de una presunción acerca de los precios del año siguiente, sin relacionarlo necesariamente con la presunción acerca de la conducta de los demás granjeros. Sin embargo, imputo a cada granjero la presunción de que los demás granjeros actuarán de cierta manera, porque la expectativa racional de un precio debe basarse en algunas expectativas acerca de lo que otros harán. Si presumimos, al igual que en el modelo más simple de la telaraña, que cada granjero actúa a base de presumir que los precios el año siguiente van a ser los mismos que los actuales, esta presunción puede ser generada por la suposición de que todos los granjeros van a producir exactamente las mismas cantidades que en el presente año, pero igualmente puede ser una presunción *faute de mieux*, un punto focal conveniente para una elección que debe realizarse. Otra posibilidad es que cada granjero actúe sobre alguna vaga presunción de que en general las cosas van a permanecer más o menos igual, sin tener una idea clara de qué cosas en particular harían la situación diferente, si *no* permanecieran igual.

Las tres observaciones restantes están orientadas a precisar más mi restricción deliberada de la noción de consecuencias no buscadas. En segundo lugar, limito el término "consecuencias no buscadas" a las consecuencias que surgen *en lugar de* las intencionadas y, por lo tanto, excluyen aquellas que surgen *además de* los resultados intencionados. No tengo interés en los autos que producen polución, sino más bien en la maquinaria anticontaminante que la genera. En tercer lugar, manejo las consecuencias no buscadas que tienen impacto en los actores mismos y no en las consecuencias no buscadas que sólo otras personas sufren. No me preocupan las fábricas que contaminan el aire de los consumidores, sino las que contaminan el aire de ellas mismas, usando óxido, etcétera. Yo creo que el grado de coincidencia o inclusión de los actores y las víctimas es una variable importante en muchos contextos, pero no sólo por simplicidad me limito al caso en el que el grupo de actores está totalmente incluido en el (y a veces es idéntico al) grupo de las víctimas. La tarea básica del científico social tiene que ser explicar *cómo reaccionan las personas frente a las situaciones creadas por ellas mismas*. Si se acepta esto, se imponen limitaciones, *a priori*, a la elección de objeto que haga el científico. Como se observa en el último apartado del capítulo 4, no tendría que contentarse con escurrir el bulto sino, más bien, buscar el objeto que esté pasando por un proceso endógeno de cambio. Efectivamente, no le será posible circunscribir un objeto totalmente autónomo que pueda ser estudiado independientemente de *todos* los factores, influencias o parámetros exógenos, pero tendría que luchar por una autonomía *local máxima*, en el sentido de que

tanto un pequeño incremento o una pequeña reducción del campo de acción haría el objeto menos receptivo de un estudio inmanentemente causal.

Por fin, quisiera enfatizar la diferencia crucial entre contrafinalidad y esas consecuencias negativas no buscadas que surgen de una sola acción. Cuando, por ejemplo, Adam Smith regañó a los mercantilistas por actuar "en contra del mismo fin que se proponen",[34] no tenía presente la contrafinalidad que surge de la suma estadística de múltiples acciones individuales (y racionales), sino sólo el hecho de que una política económica desinformada podía traer consigo resultados contrarios a los planificados. Por "desinformada" me refiero aquí a la carencia de visión de la relación de fines y medios, mientras que la contrafinalidad concuerda con una racionalidad técnica correcta, pero con presunciones incorrectas acerca de las condiciones iniciales. Para ilustrar la diferencia podemos acudir a dos estudios recientes de economía urbana. El trabajo de Jay Forrester sobre los "efectos contrarios a la intuición" de la política urbana trata de las consecuencias no buscadas que surgen de una sola decisión de planificación (tal como la decisión de crear nuevos empleos para los desempleados) cuando no se toman en cuenta los efectos indirectos y a largo plazo sobre el sistema urbano.[35] En una vena muy diferente, Thomas Schelling ha mostrado cómo puede surgir la segregación racial en una ciudad no por la acción equivocada (técnicamente desinformada) de las autoridades cívicas, sino cómo un subproducto no buscado de innumerables decisiones individuales, siendo un caso típico "que todos viven en un barrio en el que su color predomina en una proporción media algo mayor que cinco a uno", como resultado de "individuos *que buscan* una proporción no menor que cinco a cuatro".[36] No es necesario decir que los dos tipos de consecuencias no buscadas pueden producirse simultáneamente, para reforzarse o posiblemente anularse entre sí. Sin embargo, no exploraré este tema.

A la mayor parte de los ejemplos en el siguiente catálogo de casos se les dará sólo un tratamiento breve, dejando tan sólo dos para una discusión más prolongada: el modelo de la telaraña y la teoría marxista de la tasa decreciente de ganancia. Cierro esta sección con algunos comentarios acerca de lo que me propongo llamar "la falacia estructuralista en las ciencias sociales" y una falacia bastante relacionada, que se encuentra en la sociología funcionalista.

(a) En una sala de conferencias en donde todos se ponen de pie para ver mejor al orador, nadie lo logra. Paul Sweezy[37] utiliza este ejemplo como paradigma de la lógica del capitalismo, pero tendría que ser claro —y se demostrará con ejemplos enseguida— que la contrafinalidad es un mecanismo más general y no confinado únicamente a un solo modo de producción.

(b) Cuando todos desean depositar su dinero o retirar sus depósitos al mismo tiempo, ninguno logrará lo que desea (intereses y su dinero, respectivamente).

(c) Jean-Paul Sartre toma la erosión como su caso paradigmático de contrafinalidad.[38] En su ejemplo, cada campesino chino individualmente busca obtener más tierra talando árboles, pero una deforestación general induce la erosión y *menos* tierra para cada campesino que al inicio. Otros casos en los que la contrafinalidad pasa por el "campo práctico inerte" (el término de Sartre para estructuras materiales) incluyen la sobreexplotación de recursos, la tragedia de las comunas, la polución y la extinción de las especies.

(d) El crecimiento de la población y la pobreza están articulados a través de un mecanismo de contrafinalidad. En muchas sociedades campesinas, anteriores y actuales, un gran número de niños ha representado una ventaja para el individuo como una seguridad contra la vejez, pero en total la adopción generalizada de familias grandes ha tenido el efecto de reducir el ingreso *per cápita*.[39]

(e) Muchos sistemas sociales se han basado en la exportación de mano de obra o de capital; esto fue así en los antiguos estados griegos, en los estados esclavistas del sur norteamericano y en Inglaterra, en el siglo xix. Evidentemente, ésta no es una solución que pueda ser generalizada para todos los sistemas. Moses Finley[40] la llama "recurso momentáneo"; Eugene Genovese ve la imposibilidad de que todos los esclavistas sean vendedores de esclavos simultáneamente como la "contradicción más importante" en la última etapa del sistema esclavista;[41] Hegel, en unos maravillosos párrafos de la *Filosofía del derecho*[42], señalaba que un sistema de Estados, cada uno de los cuales dependa de otros como mercados para el exceso de producción, tiene que dar surgimiento a guerras imperialistas y finalmente a un derrumbe. La emigración de mano de obra, la exportación de bienes excedentes y la explotación de las colonias son, todas ellas, juegos suma cero: por lo tanto, si el despegue hacia el crecimiento industrial depende de algunos o todos estos factores, habrá algunas naciones que nunca lo lograrán o—como lo sugieren Veblen y Gerschenkron— aquellas que llegan tarde al proceso de industrialización podrán disfrutar las ventajas del atraso, que ofrece una salida de la trampa.

(f) En el sistema político también puede surgir la contrafinalidad. Podríamos presumir que, por ejemplo, algunos votantes que usualmente votan por un partido dado se sienten tan seguros de que su partido va a ganar que se permiten el lujo —de otra manera impensable— de votar por otro partido para dar a su partido un viraje a la izquierda o la derecha. Si hay un número suficiente de personas actuando de esta manera y cada votante sofisticado cuenta con la sinceridad de todos los demás, el resultado puede ser la contrafinalidad. El impacto de las encuestas de opinión sobre la votación es importante en este contexto. Las encuestas tienen un impacto bien documentado sobre la votación y podemos presumir que el impacto está mediado por la presunción de cada votante de ser él el único lector de las predicciones, o por lo menos, el único que cambia el voto por haber leído las predicciones. Herbert Simon ha demostrado que dadas determinadas condiciones es posible para quienes predicen tener en cuenta el

efecto de la predicción de manera que las encuestas sean confirmadas por la votación misma, pero, quizás, estas condiciones no son muy creíbles.[43]

(g) Un ejemplo más complejo es el muy conocido ciclo de la telaraña en la teoría económica. Para empezar, señalemos las diferencias entre las dos formas que la contrafinalidad puede asumir en la agricultura. Generalmente se supone que en los sistemas económicos precapitalistas la oferta de la producción agraria sigue una curva al menos parcialmente descendente: los campesinos pueden preferir el ocio al mayor consumo y, de todas maneras, los miembros de una familia no pueden trabajar más de 24 de las 24 horas del día. Esto puede generar contrafinalidad, de la manera siguiente: un año inicial con precios por debajo del equilibrio hace que todos los campesinos aumenten su producción para el año siguiente, presumiendo que los precios se van a mantener en un nivel igualmente bajo. Haciendo lo anterior, sin embargo, hacen bajar los precios aún más, así que tendrán que producir mayor cantidad el año siguiente y así sucesivamente en un proceso acumulativo de precios bajos y mayores volúmenes. Contrariamente, un sobre equilibrio inicial en los precios generará un movimiento acumulativo de precios más altos y menor volumen de producción.[44]

Con la curva de tendencia ascendente de oferta típica de la agricultura capitalista, la contrafinalidad adquiere la forma de un ciclo en forma de telaraña. Supóngase que la oferta y la demanda son funciones lineales del precio: $S = -a + bp, D = c - dp$, todas constantes positivas. Entonces es sencillo demostrar, e intuitivamente obvio, que si el precio inicial difiere del precio equilibrio (obtenido estableciendo $S = D$), los precios y el volumen cambiarán de acuerdo con un patrón cíclico. Los precios iniciales por encima del equilibrio generarán volúmenes por encima del equilibrio para el año siguiente, siempre suponiendo que los granjeros presuman precios constantes de un año para el otro; esto lleva a precios bajos y, para el año siguiente, volúmenes por debajo del equilibrio y precios por encima del equilibrio y así sucesivamente en un proceso que, por lo menos matemáticamente, puede darse interminablemente. El movimiento cíclico puede verse deprimido, constantemente repetido o ser explosivo, si b es menor, igual o mayor que d. En el primer caso el equilibrio es estable; en el tercero, inestable; en el segundo caso es neutral pero estructuralmente inestable.[45] Si el movimiento es depresivo, la contrafinalidad desaparecerá después de algún tiempo, si no completamente, por lo menos lo suficientemente como para que las consecuencias no buscadas se confundan con el ruido que impregna cualquier sistema social. Con un movimiento explosivo podemos presumir que después de algún tiempo los granjeros entenderán que las fluctuaciones cíclicas no se deben a los caprichos del clima ni a la intervención del gobierno, sino a sus propias acciones. Lograrán *aprender* y modificar su comportamiento.

La lógica de este proceso de aprendizaje es importante para nuestro análisis. El análisis de manual habitual[46] supone que en un primer nivel de aprendizaje cada granjero cree que los precios del año siguiente van a ser

un promedio ponderado de los del presente año y del anterior, de manera que si los precios han subido de t - 1 a t, se espera que caerán de t a t + 1 y viceversa. Sin embargo, estas expectativas sobre los precios van a generar acciones que las invaliden de tal manera que se produzca un nuevo ciclo. Este ciclo tiene más posibilidades de depresión que en el caso original; incluso con $b > d$ las oscilaciones pueden desaparecer.[47] Si no desaparecen o si tardan mucho en hacerlo, puede darse un segundo nivel de aprendizaje, de manera que los agricultores también tomen en cuenta el nivel de los precios del año anterior al pasado. Este segundo nivel de aprendizaje, a su turno, puede o no ser exitoso; en el último caso los granjeros pueden ir a un tercer nivel y así sucesivamente en un proceso que, por lo menos matemáticamente, puede continuar indefinidamente. En los casos presentes, por consiguiente, ocurrirá una de dos cosas: o bien el proceso se detiene en algún nivel o los granjeros se darán cuenta de que este modo de aprendizaje es inherentemente contradictorio. Pueden llegar a aprender que ésta no es la forma de aprender. En este caso habrán dado el paso decisivo hacia una situación *estratégica* en la cual nadie cree que los precios cambian de manera misteriosa y autónoma.

No estoy muy satisfecho con esta versión de manual acerca del proceso de aprendizaje, porque no está basada en un análisis del cambio de expectativas acerca del comportamiento de los demás. Suponer que los precios se moverán en dirección opuesta a como lo han hecho recientemente, sin articular esta presunción a alguna idea acerca de *por qué* deberían hacerlo, es un ejemplo de fetichismo en verdad frecuente, pero no muy racional. Un caso de aprendizaje "racional" (aunque contradictorio) sería el siguiente. Si el ciclo simple de la telaraña es explosivo, o no suficientemente desalentador, cada granjero puede llegar a comprender que todos están actuando bajo la presunción de que los precios permanecerán iguales de un año a otro. Asumirá entonces que los otros continuarán actuando bajo esta presunción, de manera que él —como el único— puede ajustarse de manera óptima a los precios que predominen. Sin embargo, cuando todos hacen esto, inmediatamente tenemos una nueva contradicción, puesto que *cada granjero no puede ser el único en haber aprendido*. De hecho, el proceso resultante se modelará a sí mismo cada dos años, sobre la base del ciclo original de la telaraña, de manera que obtenemos un movimiento acumulativo ascendente o descendente (como en el caso precapitalista) en vez de un ciclo. En el año t cada granjero actúa sobre la base de la presunción de que los precios t + 1 serán lo que hubieran sido si todos hubieran actuado sobre la base de la presunción de que permanecerían iguales a los del año t, y el resultado inmediato de esto será que los precios del año t + 1 del primer nivel de aprendizaje coincidan con los del año t + 2 del ciclo original. En el segundo nivel de aprendizaje cada granjero supondrá que él —como el único— es el que ha entendido lo que ocurre en el primer nivel, y así sucesivamente en un movimiento que rápidamente deviene complejo. O bien el proceso se detiene en algún nivel o los granjeros hacen la transición

al modo estratégico; ellos llegan a aprender que ésta no es la forma de aprender.

Aquí puede hacerse una breve mención de otro caso en el cual también se da una jerarquía de creencias inconsistentes, por ejemplo, en el análisis de duopolio. En la llamada"solución de Cournot" al problema del duopolio cada duopolista se empeña en maximizar las ganancias sobre la base de la presunción de que el otro va a mantener la producción constante, a su nivel real. Según Alan Coddington,[48] la norma "mantener la producción en el nivel actual" puede llamarse regla A y la norma "maximizar las ganancias sobre la base de la presunción de que el otro duopolista adoptará la llamada regla A" puede ser llamada regla B. Queda sólo un corto paso para definir la regla C como "maximizar las ganancias base de la presunción de que el otro duopolista adoptará la regla B"; cuando ambos adoptan la regla C, el resultado es el llamado "desequilibrio de Stackelberg". En ambos casos, la contrafinalidad será, por consiguiente, una característica prominente. En principio podríamos continuar definiendo las reglas D, E, F, etc; en principio, podríamos también preguntar lo que ocurriría cuando los duopolistas finalmente comprendieran que la utilización de dichas reglas conduce a la presunción bilateral de una ventaja unilateral y que por lo tanto es contradictoria.Sin embargo, el problema es que el juego correspondiente no tiene solución en una modalidad no cooperativa. Nuevamente, véanse mayores comentarios sobre el particular más adelante.

(h) Marx atribuyó la (supuesta) caída de la tasa de ganancia a las medidas (invención de maquinaria sustitutiva de mano de obra) tomadas justamente para contrarrestar esa tendencia. Argumentaba que cuando la economía avanzara hasta llegar a una tasa mayor que la del crecimiento de la fuerza de trabajo, se llegaría a un punto en el cual todo el ejército de reserva de desempleados sería absorbido por la producción y la escasez de trabajo y los salarios crecientes se establecerían. Sostuvo además que para el capitalista *individual*, la maquinaria sustitutiva de mano de obra es la respuesta racional a la caída de las ganancias causada por el aumento de los salarios,[49] pero que, cuando *todos* los capitalistas reducen la mano de obra, sobreviene una reducción de la plusvalía agregada y una caída de la tasa de ganancia. Para cada capitalista, el obrero es un factor de producción costoso, sujeto a las normas habituales de reducción de costos, en especial cuando hay escasez, pero para el conjunto de todos los capitalistas los obreros son la fuente única y última de plusvalía y, por lo tanto, de ganancias. De esta manera, el argumento tiene dos aspectos: un análisis motivacional de las acciones del empresario individual y un análisis agregado del efecto total producido cuando todos los empresarios actúan simultáneamente por el mismo motivo. Ambos aspectos de la argumentación son falaces. Sin embargo, las falacias son falacias *interesantes* y nos ayudarán a resaltar algunos puntos de vista lógicos importantes.

El análisis motivacional es inadecuado porque cuando el capitalista individual se enfrenta con altos salarios tendrá seguramente un incentivo para innovar de manera de reducir los costos, pero no un incentivo especial

para buscar innovaciones con una propensión a la economía de trabajo, porque en la competencia perfecta (que es sobre lo que escribió Marx) la demanda de trabajo del empresario individual no puede influir sobre el nivel salarial.[50] Si todos los capitalistas en busca de innovaciones se inclinaran de hecho por maquinaria sustitutiva de mano de obra, esto los beneficiaría debido a su impacto sobre la mayor demanda laboral y, por consiguiente, sobre el nivel de los salarios, pero en un mundo dominado por tomadores paradigmáticos de precios (véase el apéndice 2 más adelante) rige el axioma de que no se emprenderá una acción cuya eficacia depende de que los demás actores hagan lo mismo. Sin duda es verdad, como enfatizan tanto Marx como Neil Smelser,[51] que los capitalistas fueron inducidos a introducir maquinaria que pudiera reemplazar especialmente a los grupos recalcitrantes o combativos de trabajadores, pero no es un mecanismo generalizado. Quizás el intento de William Fellner de rehabilitar la articulación entre salarios crecientes e innovaciones en maquinaria sustitutiva de mano de obra puede verse como una generalización del argumento de Marx-Smelser. Fellner ha propuesto una lógica que podría inducir al capitalista individual a modificar su búsqueda, por ejemplo, la expectativa de que los salarios continuarán subiendo en el futuro, en cuyo caso es racional apropiarse del crecimiento futuro, utilizando menos trabajo. "En algunos casos puede desarrollarse una preferencia por las invenciones que son particularmente ahorrativas de factores de los recursos que se están volviendo más escasos, porque un *proceso de aprendizaje* puede inducir a firmas atomizadas o muy pequeñas a comportarse como si fueran suficientemente grandes para darse cuenta de que *macroeconómicamente* los factores de producción *no* son infinitamente inagotables."[52]

Tenemos que preguntar, sin embargo, por qué el capitalista individual debería tener expectativas acerca de que los salarios continúen subiendo. Tales expectativas pueden realmente ser una extrapolación de evidencia anterior,[53] pero si van a tener una base racional ésta tiene que basarse en la suposición de que otros capitalistas no van a distorsionar su búsqueda. Esto quiere decir que cada capitalista le da a su búsqueda una inclinación hacia la sustitución de mano de obra con base en la suposición de que todos los demás capitalistas no lo van a hacer: un caso claro de sistemas de creencias contradictorias, que concluye en la consecuencia no buscada de que los salarios bajan en lugar de subir. Esto se parece a un matrimonio entre las contradicciones y la mano invisible, pero en verdad, el caso es más complejo, como lo reconoce también Paul David.[54] En el próximo período los capitalistas perderán su fe en el aumento de los salarios y por ello no inclinarán sus innovaciones en determinada dirección; entonces los salarios aumentarán, y así sucesivamente, de acuerdo con el ciclo de la telaraña. Sin embargo, es diferente al de la telaraña y más relacionado con el problema del duopolio, por el hecho de que el juego que emprende no tiene solución. Estableciendo xy para la situación "Yo adopto la estrategia x y todos los demás adoptan y", en la que las variables se mueven entre el conjunto (C,D) o (Cooperar, Desertar), la situación anterior la

describen las siguientes estructuras de preferencia (comunes a todos los capitalistas): *(D,C) P (C,C) P (C,D) P (D,D)*, en las cuales la deserción se interpreta como la falta de propensión y la cooperación como una propensión al ahorro de trabajo en la búsqueda de innovaciones. Esta es la estructura del juego sin solución de la Gallina (véase el apéndice 1 del capítulo 5). Por lo tanto, si los capitalistas son tomadores de precios, no tendrán una propensión con respecto a sus innovaciones; si aprenden, su aprendizaje se hará de manera contradictoria; si comprenden que su aprendizaje encierra esa modalidad contradictoria no convergerán hacia un equilibrio (no cooperativo).

Para poder ver el error o los errores en la segunda parte del argumento de Marx es necesario tratar con algún detalle la articulación que existe entre las invenciones capitalistas para economizar trabajo, los niveles salariales, la composición orgánica del capital y la tasa de ganancia. Para facilitar la discusión, aceptemos que un capitalista empleará técnicas para economizar trabajo frente al aumento de salarios. A partir de esta presunción, la conclusión de Marx sobre la tasa decreciente de ganancia supone una argumentación de dos pasos. Primero, él supone que el promedio de la tasa de ganancia (en el conjunto de la economía) debe ser dada por la siguiente fórmula

$$r = \frac{\sum_i S_i}{\sum_i C_i + \sum_i V_i} = \frac{S}{V} = \frac{S/V}{(C/V) + 1'}$$

en la que S es la plusvalía, C el capital constante y V el capital variable, la i infrascrita indica el sector i de la economía y los demás símbolos no infrascritos representan las cantidades agregadas. (Para simplificar, se presume que todo el capital constante se renueva anualmente.) A partir de la última expresión acerca de la tasa de ganancia es evidente que si la composición orgánica C/V aumenta mientras la tasa de explotación permanece constante, entonces la tasa de ganancia tiene que caer. El segundo paso en el argumento de Marx es suponer que las invenciones para economizar trabajo equivalen a la creciente composición orgánica del capital. A partir de ello se deduce que con una tasa constante de explotación, la propensión hacia la aplicación de las innovaciones para economizar trabajo debe implicar una caída en la tasa de ganancia.

Ambos pasos en este argumento son falaces. Veremos más adelante que la tasa de ganancia tiene que estar determinada simultáneamente por los precios relativos, por lo que la fórmula de Marx no produce el resultado correcto. (Esto está relacionado con el error, más conocido, de la transformación de los valores en precios,[55] el intento de Marx de derivar los precios

tomando la suma del capital constante y el variable (medido en el tiempo de trabajo) multiplicado por $(1 + r)$. Esto es incorrecto, porque los precios tienen que igualar los *costos* (en términos monetarios) en el tiempo $(1 + r)$, puesto que los capitalistas calculan las ganancias sobre la base de los resultados monetarios más bien que sobre los valores de los factores de producción, que les son desconocidos.) Me concentraré, sin embargo, en el segundo paso y en su falacia subyacente. Al inicio podemos observar que es más bien extraña la presunción de una tasa constante de explotación, porque —dado el progreso técnico— implica, también, el aumento del salario real, la cuestión misma que la propensión a la economía de trabajo supuestamente debía prevenir. Primero argumentaré que, dada la constante del salario real (y una creciente tasa de explotación), el segundo paso en la argumentación no es válido. Las invenciones para economizar trabajo *ex ante* pueden producir una baja *ex post* en la composición orgánica. (Enseguida se clarificará lo que esto significa.) Avanzaré en la demostración de que la misma posibilidad se logra con una tasa constante de explotación.

Suponemos que cada uno de los sectores n de la economía produce un bien homogéneo con coeficientes fijos de producción, de manera que para producir una unidad de bien j, necesitamos a_{ij} unidades de bienes i y a_{oj} unidades de trabajo. Más aun, establecemos b_i para la cantidad de bienes i que entran en la canasta de consumo, que representa los salarios de trabajo por unidad temporal de trabajo. Esto implica coeficientes fijos de consumo. El valor x_j (medido en tiempo de trabajo) está determinado, entonces, por el siguiente conjunto de ecuaciones:

$$x_j = \sum_i a_{ij} x_i + a_{oj}, \qquad j = 1,2...n.$$

Los precios p_j y la tasa de ganancia r están determinados simultáneamente por otro conjunto de ecuaciones:

$$\left(\sum_i a_{ij} p_i + \sum_i b_i a_{oj} p_i \right) (1 + r) = p_j, \qquad j = 1,2...n.$$

En el último conjunto de ecuaciones establecemos $p_1 = 1$ tomando el primer bien como un numerario. Cada ecuación en el primer conjunto dice qué valor entrante debe igualar al valor saliente, cada ecuación en el segundo conjunto dice qué costos + ganancias tiene que igualar a los precios. Las condiciones para que estas ecuaciones tengan soluciones económicas significativas son bien conocidas (alguna variante de las condiciones de Hawkins-Simon).[56] No creo que, en verdad, necesitemos los valores x_j para nuestro objetivo, porque estamos tratando con las opciones que confronta el empresario individual, para quien los precios sólo tienen una realidad subjetiva. Mi definición de la composición orgánica del capital estará, entonces, formada en términos de precios más bien que en los términos usuales de valor:

$$O_j = \frac{\sum_i a_{ij} p_i}{\sum_i b_i a_{oj} p_i},$$

que es la proporción entre el costo de las maquinarias y el gasto en salarios. Suponemos que las empresas sólo adoptan innovaciones para economizar trabajo y que en cada sector hay una, que es la avanzada innovadora, mientras las otras, la siguen después de un tiempo. La empresa innovadora en el sector j tiene, por lo tanto, una posibilidad de producción $(a'_{1j}, a'_{2j}...a'_{nj}, a'_{oj})$, con las propiedades

$$a'_{ij} \leqslant a_{ij} \text{ para todas las } i,$$

$$O'_j = \frac{\sum_i a'_{ij} p_i}{\sum_i b_i a'_{oj} p_i} > O_j$$

La primera propiedad está indicada en la presunción de los coeficientes fijos de producción. Si una de las desigualdades ≤ fuera reemplazada por >, la elección entre la tecnología vieja y la nueva podría decidirse únicamente sobre bases técnicas para no ser ambigua. Los precios tendrían que ser tomados en cuenta, lo cual significa que sería necesario abandonar la afirmación central de la economía marxista, de que los valores pueden ser determinados independientemente de los precios.[57] La segunda propiedad afirma que la composición orgánica del capital en la nueva tecnología es mayor que en la vieja tecnología, *cuando ambas se calculan en base a los precios viejos*. La relevancia de esta condición debería estar clara. Si vamos a definir las innovaciones para economizar trabajo en términos relacionados con los empresarios individuales y sus motivaciones, debemos hacerlo en términos de los precios que confrontan en el momento de la elección. Por lo tanto O'_j es la composición orgánica *ex-ante* en la nueva tecnología y $O'_j > O_j$ es la condición para una propensión a la reducción de mano de obra. Por otra parte, cuando todas las empresas de todos los sectores hayan adoptado los nuevos métodos, habrá un nuevo conjunto de precios p'_j dado por

$$\left(\sum_i a'_{ij} p'_i + \sum_i b_i a'_{oj} p'_i \right) (1+r') = p'_j, \qquad j = 1,2...n,$$

y una nueva composición orgánica

$$O''_j = \frac{\sum_i a'_{ij} p'_i}{\sum_i b_i a'_{oj} p'_i},$$

para cada sector. Ahora, mi punto de vista es que bien podemos tener $O'_j > O_j > O''_j$ para todas las j, como puede ser demostrado con un simple

ejemplo bisectorial. En verdad, estableciendo $b_1 = 0, b_2 = \frac{1}{6}, a_{11} = \frac{1}{2}, a_{21} = 0, a_{01} = 1, a_{12} = \frac{1}{2}, a_{22} = 0, a_{02} = 2, a'_{11} = \frac{1}{8}, a'_{21} = 0, a'_{01} = \frac{1}{5}, a'_{12} = \frac{1}{2}, a'_{22} = 0, a'_{02} = \frac{1}{4}$, tenemos aproximadamente $O_1 = 2{\cdot}3$, $O_2 = 1{\cdot}15$, $O'_1 = 2{\cdot}9$, $O'_2 = 1{\cdot}3$, $O'_1 = 0{\cdot}5$, $O''_2 = 0{\cdot}2$. Por lo tanto, no es válida la conclusión de un incremento en la composición orgánica *ex post* a partir de una tendencia a economizar trabajo *ex-ante*. Los puristas podrían objetar que Marx en sus referencias al cambio en la composición orgánica del capital siempre habla en términos de valores. Siendo $\bar{O}$ la composición orgánica en términos de valores, un contraejemplo tendría que tener la siguiente estructura: $O'_j > O_j$ y $\bar{O}_j > \bar{O}''_j$. Un coeficiente igual al establecido arriba $\bar{O}_1 = 6, \bar{O}_2 = 3, \bar{O}''_1 = \frac{40}{87}, \bar{O}''_2 = \frac{128}{609}$. Si ahora volvemos al caso de la tasa constante de explotación, esto implica que entre los ajustes que deben hacerse a la nueva tecnología hay un cambio en b_i. Primero, tenemos que encontrar un conjunto b'_i tal que $\sum_i b_i x_i = \sum_i b'_i x'_i$ en la que x'_i sea una unidad de bien $_i$ en la nueva tecnología y, segundo, debemos calcular el nuevo conjunto de precios p''_i desde las ecuaciones

$$\left(\sum_i a'_{ij} p''_i + \sum_i b'_i a'_{oj} p''_i \right) (1 + r'') = p''_i, \qquad j = 1,2...n,$$

a partir de las cuales pueden calcularse las composiciones orgánicas. Una vez más, el mismo ejemplo sirve para ser utilizado en la demostración de que las innovaciones para economizar trabajo pueden darse conjuntamente con la caída de las composiciones orgánicas *ex post*, aun con una tasa constante de explotación.

La razón por la que ésta es una falacia *interesante* es que Marx olvida los efectos mismos de la interacción a los cuales él atribuye la contrafinalidad del capitalismo. De acuerdo a Marx, las crisis del capitalismo se deben a que cada individuo toma sus propias decisiones sobre la base de un equilibrio parcial, asumiendo equivocadamente que todos los demás actores deben ser tomados como parámetros para resolver sus decisiones. En su razonamiento, Marx mismo comete el error de llegar a conclusiones partiendo del equilibrio parcial para llegar al general: comete la falacia de composición cuando olvida la diferenciación entre la composición orgánica de la nueva tecnología cuando se calcula en viejos o nuevos precios. A partir del hecho de que una innovación dada economiza trabajo cuando otras cosas se mantienen constantes, Marx considera ilegítimamente que esto sigue siendo verdadero cuando todas las innovaciones son consideradas simultáneamente. Además de la falacia de composición podemos hablar de la falacia patética: la contradicción que Marx discierne en la economía capitalista no es más que la proyección de un error en su propio razonamiento.

La falacia es más sorprendente porque Marx sabía muy bien cuál es la importancia de la calibración simultánea en la economía. En verdad creo que su reconocimiento de esta idea es el primero en la historia del pensamiento económico y un paso adelante decisivo en la visión "histórica"

del valor.[58] Sin embargo, Marx era demasiado hegeliano como para liberarse completamente del modo de pensar histórico. Para ser más preciso, creo que la teoría hegeliana del Espíritu objetivo —la mano muerta del pasado que se encarna en las instituciones del presente— tal vez esté en la raíz de la teoría falaz de la tasa decreciente de la ganancia. Desde el punto de vista subjetivo del obrero, los bienes de capital lo van dominando cada vez más, hasta transformarlo en un simple apéndice de la máquina.[59] En el desarrollo del pensamiento de Marx, el paso psicológico desde esta dominación cualitativa del trabajo por el capital a una dominación cuantitativa (es decir, un incremento de la composición orgánica del capital) puede haber sido más corto que el paso lógico. El incremento en el *volumen físico* (y en la importancia subjetivamente sentida) de "máquinas por trabajador" puede haber obstaculizado el reconocimiento de un posible descenso en el *valor* del capital constante por unidad variable de capital.

Concluyo esta sección con unas cuantas observaciones acerca de lo que propongo llamar *la falacia estructuralista en las ciencias sociales*, una falacia que, de manera injustificada, equipara la relación entre motivaciones, acciones y consecuencias en el análisis sociológico con la correspondiente relación en la racionalidad individual.

Veamos las siguientes modalidades de inferencias en las que las acciones x son posibles y todos los actores son racionales:

(1) Cuando el actor a hace x, esto produce el conocido y deseado efecto y. Por lo tanto, a hará x.

(2) El actor a hizo x.

Cuando el actor a hace x, esto produce el conocido y deseado efecto y.

Por lo tanto, el actor a hizo x para obtener y.

(3) Cuando todos los miembros del grupo A hacen x, esto produce el conocido y deseado efecto y.

Por lo tanto, todos los miembros de A harán x.

(4) Todos los miembros de A hicieron x.

Cuando todos los miembros de A hacen x, esto produce el conocido y deseado efecto y.

Por lo tanto,los miembros de A hicieron x para obtener y.

Las primeras dos de estas inferencias pueden suponerse válidas por la definición de actor racional. Alguien podría cuestionar la segunda inferencia sobre la base de que podemos hacer x para obtener z, incluso si prevemos que esto también tendrá el efecto beneficioso y. Esto es más parecido a una sobredeterminación causal, como la presencia simultánea de dos determinaciones causales, y y z, cada una de los cuales es suficiente para causar x. Sostengo que, si alguien dijera "Yo hago x para obtener z, pero preveo que también obtendré y, y si pensara que x no me traería z, sin embargo hago x para obtener y", entonces él realmente hace x para obtener y,

de manera que la segunda inferencia está justificada. La tercera inferencia que ha sido llamada falacia anárquica y falacia liberal,[60] es obviamente equivocada, como se ve apelando al Dilema del Prisionero. Las personas, con frecuencia, no actúan colectivamente de manera óptima, pero es tentador concluir que cuando *sí* actúan de manera óptima, lo hacen para obtener un resultado óptimo. Esta conclusión es la falacia estructuralista, por ejemplo, la cuarta inferencia anteriormente vista. Algunos ejemplos justificarán, así lo espero, que adicionemos otra falacia a una larga lista de modos inválidos de inferencia.

El primer ejemplo es la cuestión de Marx-Fellner, a la que acabamos de referirnos, la cual surge si representamos como "*x*" la "búsqueda de innovaciones para economizar trabajo" e "*y*" la "reducción del nivel de salarios". Un segundo ejemplo es dado por Eric Hobsbawm cuando alerta contra la explicación falaz de los orígenes del capitalismo que convierte a las economías externas en motivaciones para los empresarios individuales.

> El mismo proceso que reorganizó la división social del trabajo aumentó la proporción de los trabajadores no agrícolas, diferenció al campesinado y a las clases asalariadas existentes, también creó hombres que necesitaban comprar en dinero para satisfacer sus necesidades: clientes para bienes. *Pero éste es el punto de vista del analista, no del empresario, quien decide si revolucionar o no su producción.*[61]

Un poco más complejo es el siguiente ejemplo. Aquí la falacia es cometida por E.P.Thompson en un importante artículo sobre las revueltas campesinas en la Inglaterra preindustrial. Afirma que había revueltas recurrentes y observa que, aunque fracasaban en sus objetivos inmediatos, a largo plazo obligaban a las clases propietarias a comportarse de manera más moderada. De estas dos proposiciones parece concluir que ese éxito a largo plazo da una explicación *motivacional* de las revueltas. Así interpreto su pregunta retórica sobre si "las revueltas hubieran continuado durante muchas décadas, en verdad, centenas de años, si hubieran fracasado consistentemente en lograr sus objetivos".[62] Esta explicación, que no es imposible, presume un grado más bien improbable de solidaridad intergeneracional. Un enfoque más económico sería explicar la persistencia de las revueltas por la persistencia de sus causas. De todos modos, la afirmación motivacional no puede ser deducida de premisas no motivacionales. Del hecho de que cada generación se levantó y de que fue para beneficio de todos que lo hiciera, no podemos concluir que todos se levantaron para llevar a cabo este beneficio.

El último ejemplo es muy diferente, pero exhibe la misma estructura general. Lo hemos tomado de las controversias metodológicas que se agitan en las bases de la teoría evolucionista. En el trabajo más importante de la primera fase de esa discusión George Williams denunció la "concepción errada .... de que cuando se demuestra que un determinado proceso biológico produce un beneficio determinado, se ha demostrado *la* función

o por lo menos *una* función del proceso".[63] De sus ejemplos se deduce claramente que lo que Williams tiene en mente son los efectos benéficos que surgen de la interacción entre los individuos. Puede tomarse un ejemplo de un texto de Stanley Salthe:

> Otro ejemplo de una estructura supraindividual puede verse en el espaciamiento de individuos durante el anidamiento de una colonia en un criadero de garzas. Cada pareja tiene un pequeño territorio en el cual está localizado su nido. Cualquier individuo que no sea de la pareja e ingrese a ese territorio es atacado y expulsado. Las crías tampoco son bienvenidas en el territorio de ninguna pareja, incluyendo el de sus padres. Ellas ocupan espacios estrechos, entre un territorio y otro, en su mayoría ubicados en áreas relativamente seguras frente al ataque de los depredadores. Sus padres no se preocupan por protegerlos sino por expulsar a los depredadores. Si lo logran, también protegen "no intencionalmente" a cualquier cría que ocupa espacios cercanos a su territorio.[64]

En el formalismo de la inferencia (4) tenemos x = padres que defienden su territorio e y = protección de las crías: las economías externas o los efectos sobre la acción individual adaptativa x en sí mismos no pueden establecer una presión selectiva para este comportamiento. Otro ejemplo de esta falacia biológica podría ser la de los efectos no intencionales que surgen de la pleiotropía, o sea, del hecho de que un solo gen influya sobre diversas características fenotípicas. Podemos suponer que el gen *a* se expresa en las características A y B. A es una característica que confiere ventajas reproductivas al organismo de una manera directa, sin ningún efecto de interacción. B es una característica altruista, cuyo beneficio para el individuo es similar, en su estructura, a la del Dilema del Prisionero: para cada organismo lo mejor (en términos de capacidad reproductiva) es ser el único sin esa característica. Luego, seguiría como lo mejor que todos los organismos la tuvieran; luego, que ningún organismo la tuviera; y lo peor, ser el único organismo con esa característica. (Williams[65] cita el caso de los cardúmenes de peces como ejemplo posible de propiedades "egoístas"de organismos; otros ejemplos incluyen la agresión y advertencias discriminatorias.) Un gen *a* nunca podría haber sido seleccionado por causa de B solamente, por lo menos si hacemos abstracción de la posibilidad de selección de grupo. Por el contrario, cualquier organismo que exhibe B podría ser —*ceteris paribus*— un perdedor en la lucha por las ventajas reproductivas, de manera que la característica desaparecería pronto de la población. Sin embargo, si las ventajas conferidas por A fueran suficientemente amplias, la pérdida que implica B podría ser tolerable. En este caso, una característica evidentemente "benéfica", tal como el comportamiento altruista en el Dilema del Prisionero, realmente no "tiene una función" (no ha sido seleccionado para) sino que es sólo un efecto colateral "no intencional" de un gen por cuyo producto principal ha sido seleccionado. Me parece que ésta es una de las cuestiones a las que Robert Nozick se refería en su debate sobre las explicaciones de la mano invisible: lo que parece ser

adaptación intencional (o funcional) puede tan sólo ser un producto colateral de acciones (o selecciones) emprendidas para otros fines (o "fines").

El paradigma de la explicación funcional en la sociología, comprende una línea de razonamiento que, a mi entender, es muy cercana a la falacia estructuralista, siendo, sin embargo, distinta. La característica común de ambas modalidades de razonamiento es que consecuencias objetivas favorables de algún conjunto de acciones son vistas como *explicativas* de las acciones. En la falacia estructuralista estas consecuencias se transforman en *motivaciones* individuales para la acción, mientras que las explicaciones funcionalistas postulan algún tipo de *realimentación* causal, de efecto a causa. Yendo a las exposiciones corrientes de la sociología funcionalista de Robert Merton y Arthur Stinchcombe,[66] creo que un paradigma de explicación funcional sería el siguiente. Decimos que el *fenómeno X está explicado por la función Y para el grupo Z*,si y sólo si:

(1) *Y* es un efecto de *X*;
(2) *Y* es benéfico para *Z*;
(3) *Y* no es buscado por los actores que producen *X*;
(4) *Y* no es reconocido por los actores en *Z*;
(5) *Y* mantiene *X* por realimentación causal pasando a través de *Z*.

Casi nunca se reconoce que muy raramente se proponen teorías que satisfagan todos estos criterios. Con bastante frecuencia, se satisfacen los criterios (1), (2), (3) y (5) pero no el criterio (4). Podríamos hablar de *procesos de filtración* o quizás de *selección artificial*. El fenómeno en cuestión es mantenido precisamente porque algún grupo lo percibe como benéfico, aun cuando los actores que lo producen no se dan cuenta de ese efecto, y quizá, se oponen a él. Cuando el Departamento de Defensa de los Estados Unidos adjudica fondos a los científicos, cada uno de ellos los aplica en un proyecto que le parece el más interesante sobre bases puramente científicas. El fin resultante será una tendencia en la composición de la investigación, incluso si cada científico individual cree que sigue su propia tendencia individual.[67] En otros casos, los criterios (l),(2), (3) y (4) son satisfechos, mientras que no lo es el (5). Por ejemplo, Marx y Mandeville argumentaron[68] que la delincuencia tenía un impacto benéfico en la economía capitalista, aunque no fuera la intención de los delincuentes ni el hecho fuera reconocido por los legisladores y empresarios. El único caso convincente con el que me he cruzado y en el que los criterios del (1) al (5) son satisfechos simultáneamente es la explicación del comportamiento de maximización de la ganancia en empresas que sostienen que las empresas que no utilizan reglas prácticas de maximización de la ganancia van a dejar de existir, de manera que sólo las maximizadoras de ganancias quedarán.[69] Esto, en verdad, es un caso de *selección natural* operando a nivel social, pero afirmo, que dichos casos son sumamente raros.

La falacia comprendida en el razonamiento funcional es la inferencia de una realimentación causal a partir de las propiedades del (1) al (4). Por

ejemplo, si se puede demostrar que la neutralidad aparente del estado en las sociedades capitalistas modernas es más acorde con los fines capitalistas que un gobierno abiertamente sesgado, un argumento funcional marxista muy conocido tiende a concluir que este efecto benéfico *explica* la neutralidad del estado.[70] A no ser que se impute la realimentación a algunos oficios filtrantes de un comité ejecutivo de la burguesía, es difícil ver cómo opera. La realimentación tiene que ser *demostrada* y no solamente *postulada*, a diferencia del caso biológico en el cual la existencia de un efecto benéfico da una razón *prima facie* para pensar que es un producto de la selección natural. Argumenté anteriormente que es una falacia pensar que *tiene* que ser el producto de la selección natural, pero de todos modos existe un mecanismo general para la realimentación causal en la evolución. Un mecanismo como éste no existe en las sociedades, por lo que el paso de (1)— (4) al (5) no puede justificarse ni siquiera como una idea reguladora.[71] La propia falacia estructuralista podría haber explicado la neutralidad del estado imputando determinadas intenciones maquiavélicas a sus funcionarios. Frente a la creciente imposibilidad empírica de esta línea argumental, los marxistas han tendido a retirarse hacia la fortaleza impenetrable del funcionalismo, postulando o asumiendo implícitamente, alguna realimentación causal. O sea, de una posición falsada se han retirado a otra, infalsable. En verdad, Karl Popper ha demostrado que todas las afirmaciones existenciales positivas, tales como "Existe una secuencia finita de dísticos elegíacos latinos tal que si se pronuncian de manera apropiada en un determinado tiempo y lugar esto es inmediatamente seguido de la aparición del Diablo",[72] son no falsables (aunque verificables) y, por lo tanto, quedan fuera del campo de la ciencia. De manera similar, la afirmación de que existe alguna realimentación causal es no científica si no está basada en el razonamiento general, que podría prestarle alguna posibilidad *prima facie*, ni en la demostración de algún mecanismo específico.

Suboptimalidad

Propongo llamar "suboptimalidad" a la segunda variedad de contradicciones sociales: *la realización deliberada de una solución no cooperativa, que es Pareto-inferior a algún otro conjunto de resultados obtenible por elecciones individuales de la estrategia.* Es decir, cuando todos los jugadores adoptan las estrategias de solución sabiendo que los otros hacen lo mismo y que todos pueden obtener, por lo menos, lo mismo y uno de ellos más, que si alguno o todos ellos hubieran adoptado estrategias de solución diversas (los juegos carentes de una solución no cooperativa serán discutidos en la próxima sección). El ejemplo paradigmático de suboptimalidad es, por supuesto, el Dilema del Prisionero. Antes de continuar con algunos ejemplos, es necesario examinar dos problemas generales: ¿Cuál es la relación entre la contrafinalidad y la suboptimalidad y qué expresa el término "contradicción"en el caso de la suboptimalidad? La respuesta a la segunda pregunta depende, de manera importante, de la respuesta a la primera: la

suboptimalidad está subsumida bajo el encabezamiento de contradicciones porque en muchos casos no es diferenciable de la contrafinalidad.

La suboptimalidad, como se definió, es una noción de la teoría de los juegos. Presupone información perfecta y anticipaciones correctas. Sin embargo, hay que enfatizar el hecho de que la transición de una situación preestratégica a otra estratégica, con frecuencia no modifica el comportamiento de los actores. En el Dilema del Prisionero, por ejemplo, la contrafinalidad es generada cuando todos hacen x con la esperanza de ser los únicos que hacen x,y en la suboptimalidad, cuando todos hacen x por temor a ser los únicos que no hacen x. La esperanza de ser alguien que se define solo cuando se generaliza, genera contrafinalidad; el temor a ser un bobo, no. No hay diferencia en el comportamiento observado entre los actores sólo por sus expectativas. En el caso de la suboptimalidad, las consecuencias son previstas más bien que imprevistas, intencionadas más que no buscadas.[73] En otros casos, sin embargo, la transición de la racionalidad paramétrica a la estratégica (ver apéndice 2) *hace* una diferencia en el comportamiento de los actores. El modelo de la telaraña ilustra este caso. Aquí, el precio competitivo es la solución al juego, siendo éste el único punto de equilibrio. Sin embargo, ello no significa que todas las contradicciones desaparecen. Como ha sido observado por Luce y Raiffa,[74] la estrategia competitiva puede verse también como comportamiento no cooperativo en el Dilema del Prisionero, como opuesto al comportamiento cooperativo de la restricción monopólica o a la maximización conjunta de la ganancia. Con información imperfecta, los granjeros están constantemente frustrados porque sus expectativas y planes fracasan. Con información perfecta, están frustrados porque entienden que podrían haber mejorado su suerte actuando de manera diferente a la indicada por la racionalidad.

La diferencia principal entre el Dilema del Prisionero y el juego de la telaraña es que en el primero los actores tienen una *estrategia dominante*. En juegos con una estrategia dominante suele ser difícil decidir si un comportamiento dado es un caso de contrafinalidad o de suboptimalidad, a no ser que se tenga evidencia directa acerca de las expectativas de los actores. Sin embargo, si en un juego sin estrategia dominante se lleva a cabo una solución suboptimal podemos concluir de inmediato que es un caso de suboptimalidad y no de contrafinalidad, porque las estrategias de solución se escogerán sobre la presunción (correcta) de que los otros también las están escogiendo. No es fácil decir con cuánta frecuencia la solución se realiza en juegos de la vida real sin estrategias dominantes. Los estrictos requisitos de información necesarios para que se realice la solución hacen improbable a priori que los actores se arriesguen a actuar sobre la base de la presunción de que los otros también están tomando el mismo riesgo. Con el más pequeño elemento de incertidumbre, el curso normal será recurrir a una estrategia de máxima y mínima. Pueden mencionarse algunas

excepciones a esta conclusión general. Primero, el grupo puede ser tan pequeño y estable (ver también la última sección de este capítulo y el apéndice 1) que las personas lleguen realmente a conocerse y confiar mutuamente. Segundo, la información pertinente puede ser dada por un agente externo, por ejemplo, el gobierno. Si una situación dada es un Juego de Certeza mientras cada actor equivocadamente cree que para todos los demás actores es un Dilema del Prisionero, el gobierno puede entonces dar información acerca de la estructura real de preferencia que pueda llevar a los actores a converger en una solución. Esto podría concebirse como importante en casos tales como la evasión de impuestos y el esfuerzo en el trabajo. (para mayor discusión del Juego de Certeza, ver apéndice 1.) Tercero, la repetición de la situación en una forma esencialmente igual puede hacer surgir un *proceso de aprendizaje*, como en el caso de la telaraña.

Un caso especial de juegos sin estrategia dominante es aquel en el que la solución exige que el actor utilice *estrategias mixtas*. Esta noción es decisiva en parte del reciente trabajo de Raymond Boudon, que veremos un poco más adelante. Hasta lo que yo sé, no hay evidencia empírica de que la gente escoja deliberadamente estrategias mixtas en situaciones de la vida real que comprendan un número tan grande personas como en los casos de Boudon. Sería muy interesante, en verdad, si dicha evidencia pudiera encontrarse, porque —como se discutirá en el apéndice 2— éste podría ser el puente entre el vacío que existe en el presente entre las imágenes intencionales y causales del hombre en las ciencias sociales. Sin embargo, ciertamente, aún somos muy ignorantes en estas cuestiones.

Es posible, ahora, explicar por qué propongo subsumir la suboptimalidad bajo el encabezamiento general de contradicciones sociales. El argumento principal es que en muchos casos la contrafinalidad y la suboptimalidad son idénticas en cuanto a su comportamiento aparente. Un argumento adicional es el siguiente. Como intento demostrar en la próxima sección, tanto la contrafinalidad como la suboptimalidad tienden a generar acción colectiva. En el primer caso, por la tensión entre *la intención y el resultado*, en el segundo, por la tensión entre *lo real y lo posible*. Estas tensiones no son contradicciones en sí mismas, en el sentido riguroso del término utilizado aquí. Sin embargo, apuntan a la necesidad de una terminología unificada. Hay tanto una superposición formal como una afinidad substancial entre contrafinalidad y suboptimalidad. Debería, quizás, añadir que *no* deseo defender mi tesis de que las contradicciones generan acción colectiva, dándoles un nuevo nombre a todas las "contradicciones" causantes de esa acción, sino persuadir al lector de que la pequeña extensión misma expresada en el paso de la contrafinalidad a la suboptimalidad tiene la ventaja de unificar y simplificar la discusión. No tiene demasiada importancia que el lector permanezca con dudas o desconfianza, porque la sustancia del argumento se explica en lo que resta del capítulo. Procedo ahora, como antes, a utilizar algunos ejemplos diferentes para sentar puntos conceptuales que parecen ser importantes.

(a) Los ejemplos corrientes de suboptimalidad que surgen en las situaciones que tienen la forma general del Dilema del Prisionero son externalidades, los bienes públicos y las organizaciones. Como son tratadas exhaustivamente en el clásico-moderno de Mancur Olson acerca del tema, lo mismo que en muchos otros trabajos,[75] seré más bien breve al cubrir este punto. Las externalidades —tales como la polución— pueden surgir como contrafinalidad y como suboptimalidad, pero afirmo que la no provisión de bienes públicos surge, la mayor cantidad de veces, en el modo de la suboptimalidad. Pienso que típicamente las personas o bien no se darán cuenta de la existencia y la posibilidad del bien público en cuestión (la ausencia del control de la polución no fue sentida en Grecia Antigua) o entenderán que el problema del que se corta solo impedirá la realización del bien. El caso intermedio, en el cual cada actor espera ser el único que obtiene beneficios del bien sin contribuir a los costos, no parece muy factible. Es posible que el temor a ser un bobo anime más a los actores que la esperanza de ser alguien que se corta solo.[76] Este es un punto difícil y no demasiado importante. La principal idea que quiero transmitir es que no debemos hablar sobre suboptimalidad cuando las estrategias que darían el resultado Pareto-superior no pertenecen a las posibilidades subjetivas de los actores.

La suboptimalidad puede también surgir de la ausencia de organizaciones, como ocurre en el ejemplo de la telaraña. Es mejor para todos los granjeros formar un cártel para maximizar las ganancias colectivas que si todos las maximizan individualmente, pero para el granjero individual sería aun mejor romper el acuerdo del cártel e incrementar la producción para explotar los precios altos causados por la restricción a la producción hecha por el cártel. Como todos los granjeros entienden esto no se formará ningún cártel. Un argumento similar puede regir en la formación de sindicatos, tal como la analiza Mancur Olson. Incluso si los trabajadores entienden que van a estar mejor para negociar si todos se organizan que si cada uno se encuentra solo (una visión que demoró bastante tiempo en surgir, en el siglo XIX, en Europa, donde se creía que el único efecto del sindicato era redistribuir el ingreso en la clase trabajadora), el trabajador individual puede verse tentado a tener su torta y comérsela, como el que se corta solo. Durante una huelga, por ejemplo, el rompe huelgas no obtendrá el salario usual (el que tenía antes de la huelga) pero puede reclamar salarios más altos, debido a su monopolio temporal del poder laboral. Después de la huelga, por consiguiente, se beneficiará con los salarios más altos obtenidos por los huelguistas. Sin embargo, si cada uno razona de esta manera, ninguno obtendrá beneficio. En la próxima sección discutiré este problema con más detalles.

(b) Una utilización interesante de la lógica de contrafinalidad y suboptimalidad se encuentra en un trabajo de Raymond Boudon sobre la sociología de la educación. En contraste con la aproximación clásica a la educación como un bien público, que se convierte en un problema del que

se corta solo en el que todos quieren beneficiarse (por medio del incremento de la productividad) con la educación de los demás —un Dilema del Prisionero con muy poca educación como solución—, Boudon argumenta que en la sociedad industrial moderna hay una tendencia hacia una excesiva educación, ya que todos desean tener más educación que los otros para disfrutar de un ingreso diferencial:

> El famoso teorema de la telaraña sugiere que las personas pueden comportarse racionalmente (o sea, tomar decisiones aparentemente adaptadas a sus objetivos y, sin embargo, haciéndolo obtienen resultados contradictorios con esos objetivos. Esto es exactamente lo que encontramos: cada individuo tiene un beneficio definido al intentar obtener tanta educación como le sea posible: a mayor nivel de educación, más favorables son las expectativas del estatus. Pero, tan pronto como todos los individuos desean más educación, los niveles tienden a degenerarse y esto tiene el efecto de incitar a las personas a exigir más educación en el período siguiente.[77]

El problema, como fue planteado aquí, es de contrafinalidad y no de suboptimalidad, pero esto no es esencial. (La referencia al modelo de la telaraña es engañoso, porque en el modelo la diferencia entre un contexto preestratégico y uno estratégico es esencial, como ya lo vimos. El modelo de la telaraña también describe un movimiento cíclico, mientras que el modelo de demanda de educación, genera un movimiento ascendente de exigencia.) En otro trabajo, Boudon y sus colaboradores han explorado de manera más explícita y detallada la lógica de la suboptimalidad en la educación.[78] Intentando explicar por qué los estudiantes franceses no buscan (de manera numerosa, como ha sido previsto por las autoridades educativas) el camino más corto en la educación superior disponible en los Institutes Universitaires de Technologie, sino que prefieren el camino más largo de las instituciones clásicas de estudios superiores, proponen un modelo incorporando los siguientes supuestos. En un conjunto de 20, los estudiantes tienen que elegir entre el camino más corto y el más largo. Ellos saben que al final del camino corto van a encontrar un empleo por 1 franco, mientras que todo depende del número de estudiantes que lo hayan escogido. Concretamente, se presume que para los estudiantes con la educación del camino largo hay seis empleos de 2 francos y ocho de 1. Estos empleos se llenan automáticamente si hay candidatos calificados; si hay demasiados candidatos, el excedente no conseguirá ningún trabajo.

Dadas estas presunciones, es sencillo ver que la remuneración total de los estudiantes (y la esperada por cada estudiante) es maximizada si no menos de 6 estudiantes y no más de 14 toman el camino largo, de manera que 6 parece ser la capacidad mínima para el camino corto. Este razonamiento, sin embargo, pierde de vista que las cosas pueden parecer muy diferentes desde el punto de vista del estudiante individual. Si x estudiantes tomaran el camino largo ($6 \leqslant x \leqslant 14$), cada uno tendría que utilizar una estrategia mixta y escoger el camino largo con la probabilidad $x/20$ y el camino corto con la probabilidad $1\text{-}x/20$. Este conjunto de estrategias mix-

tas, desafortunadamente, no es un punto de equilibrio. De hecho, es sencillo ver que el camino largo es la *estrategia dominante*, de manera que la del camino corto no será elegida por nadie, dejando una capacidad excedente de seis. La remuneración total caerá de 26 a 20 francos y la expectativa de remuneración de 1,3 a 1, un caso claro de suboptimalidad. Esta línea de argumentación es ingeniosa e interesante, pero los resultados dependen demasiado de las presunciones numéricas específicas. Boudon y sus colaboradores afirman[79] que la conclusión materialmente no cambia si presumimos que, al final del camino corto, el estudiante tiene la certeza de obtener 1,5 francos, pero esta afirmación no es correcta. En este caso no hay estrategia dominante y la solución del juego requiere que cada estudiante elija el camino largo con probabilidad 0,6 y el camino corto con probabilidad 0,4, mientras que la combinación óptima colectiva sería (0,3,0,7), dando los resultados esperados de 1,5 y 1,65, respectivamente. Por lo tanto, *si* los estudiantes eligen las estrategias combinadas de la solución estamos ante un caso de suboptimalidad, como se define aquí; pero he argumentado que éste es un gran *si*. Más probablemente, cada estudiante elegirá su estrategia educativa en base al principio de maximin o en base a alguna presunción concerniente al comportamiento de los otros, por ejemplo, la de que otros estudiantes del grupo escogerán en la misma proporción que el grupo anterior. La última posibilidad podría generar un verdadero ciclo de telaraña, con muchos y pocos estudiantes, eligiendo alternadamente el camino corto.

En el trabajo que acabamos de discutir no hay una mención explícita a la estrategia combinada. Uno de los casos discutidos por Boudon finaliza utilizando estrategias combinadas y esto constituye una objeción a su argumento.[80] En un trabajo posterior,[81] Boudon respondió a sus críticos y aprovechó la noción de estrategias combinadas con un objetivo muy interesante: la reinterpretación de los clásicos hallazgos de *El soldado norteamericano* sobre la inesperada correlación entre las posibilidades de promoción y el nivel de frustración.[82] Imagínese un grupo de 20 PM (policía militar) y otro de 20 miembros de la fuerza aérea. El primero, es un grupo caracterizado por pocas posibilidades de promoción y bajo nivel de frustración, mientras que en el segundo, los potenciales de promoción y de frustración son altos. Supóngase que para participar en el "juego de promoción" es necesario arriesgar el equivalente a 1 franco de esfuerzo. Los promocionados reciben el equivalente a 5 francos de remuneración. Cinco de cada veinte miembros de la fuerza aérea serán promovidos, mientras que sólo dos de cada veinte PM lo lograrán. Es sencillo ver, entonces, que todos los miembros de la fuerza aérea arriesgarán 1 franco, siendo el resultado que 15 de los 20 saldrán frustrados, mientras que sólo 10 PM entrarán al juego, de los cuales sólo 8 necesariamente saldrán descontentos. El mecanismo que hace que los 10 PM entren al juego es, por consiguiente, que cada uno de los 20 utiliza la estrategia combinada de entrar al juego con una probabilidad 0,5. Mayor potencial para ser promovido conduce no sólo a una mayor cantidad de participantes en el juego de

promoción sino a una diferencia en la participación que excede la diferencia en el potencial de promoción, lo cual significa más frustración. De nuevo, sólo puedo afirmar que la validez de este elegante análisis depende de la posibilidad empírica de estrategias combinadas en las situaciones de la vida real.

(c) Los ejemplos restantes tratarán de diversas maneras sobre consideraciones dinámicas, es decir, la suboptimalidad en el tiempo. El primero de estos ejemplos es el problema de las preferencias temporales inconsistentes, discutidas en el capítulo 4, donde prometí un análisis posterior. Reafirmaré que el problema es el de un juego entre generaciones sucesivas. Supóngase en primer término que cada generación está interesada solamente en su propio consumo; por lo tanto, para todas las generaciones es mejor emprender alguna inversión que si ninguna lo hace, pero para cada generación es mejor consumir todo el producto existente mientras se beneficia con la inversión de generaciones anteriores. Para el caso de dos generaciones A y B (A precede a B), la matriz de resultados tendría la siguiente estructura ordinal:

		B	
		Inversión	Consumo
A	Inversión	10,10	10,15
	Consumo	15,5	15,8

Vemos que la estrategia de no inversión domina a la estrategia de inversión para cada generación. El "juego" no es un Dilema del Prisionero porque A no puede resultar lesionada por la represalia de B. Sin embargo, en otros casos es posible darle un sentido a tal represalia, por ejemplo, si la utilidad de una generación dada está en función no sólo de su propio consumo sino del consumo disfrutado por generaciones siguientes. Esto, por consiguiente, no implica que el comportamiento futuro pueda tener un impacto sobre la utilidad presente, sólo que las expectativas presentes sobre el comportamiento futuro pueden tener dicho impacto. Suponiendo una secuencia infinita de generaciones, el problema del equilibrio intergeneracional en la teoría de juegos[83] puede ser expuesto de la siguiente manera: ¿Existe una secuencia *Si* de proporciones ahorrativas, de manera que para todas las i sea verdadero que si la generación i espera que sus sucesores elijan proporciones $s_{i+1}, s_{i+2}...$, ella misma elegirá S_i para la maximización de utilidades?

Para ver cómo surge este problema traemos a colación que una generación "ingenua" intentaría maximizar utilidades fijando tanto la proporción de sus ahorros como la de todas las generaciones futuras. El obstáculo, como vimos en el capítulo 4, es que la siguiente generación puede querer cambiar la proporción de ahorro que se fijó para ella, de

manera que en el sentido *ex post facto* la primera generación no lograría sus expectativas. En este sentido —algo forzado— podemos decir que las generaciones posteriores pueden reaccionar contra sus predecesores o que una generación puede ser "engañada" por sus sucesores y experimentar contrafinalidad. También hemos visto que dicha revisión continua de los planes se llevará a cabo si las preferencias temporales tienen una forma no exponencial, es decir, si los valores actuales de una unidad de utilidad en los años futuros no forman una serie exponencial decreciente. Esto puede ocurrir de la siguiente manera. Supóngase que hay dos clases de descuento del futuro, uno independiente del tiempo de todas las utilidades futuras, no importa el tiempo durante el cual son experimentadas, de manera que la generación t está de acuerdo con una unidad de utilidad experimentada por la generación $t + T$ al valor actual $(1 + r_1)^{-1}$. Segundo, hay un descuento exponencial dependiente del tiempo que da una unidad de utilidad experimentada en $t + T$ al valor actual $(1 + r_2)^{-T}$. Combinando las dos, da un descuento no exponencial del futuro con la necesidad de una revisión continua de planes ingenuos.

No hay, por consiguiente, imposibilidad lógica de presumir que alguna generación pueda imponer su deseo sobre el futuro, aun si es difícil ver cómo puede realizarlo en el nivel institucional. Hay, sin embargo, una imposibilidad lógica en la presunción de que todas las generaciones puedan hacerlo si todas tienen preferencias temporales no exponenciales. Si todas intentan maximizar "ingenuamente", todas van a ser engañadas en el sentido *ex post facto*. Aquí, como en los casos examinados en la sección previa, las presunciones no universalizables acerca del comportamiento de los demás llevan a consecuencias no intenciondas. Antes de continuar la consideración acerca de la transición de la contrafinalidad a la suboptimalidad que se observa cuando las generaciones ingenuas logran un punto de vista más sofisticado, me gustaría señalar que estamos tratando aquí con un problema básico de la teoría de la planificación democrática. Francis Sejersted habla acerca de "un conflicto interno"[84] —en nuestros términos, una contradicción— en la noción de democracia. Extendiendo el campo del ejercicio actual del poder al límite, la democracia también socava la eficacia del poder. En el límite, la democracia completa significa que cada generación está libre de las obligaciones dejadas por sus predecesores, pero esto también implica que ella misma es incapaz de atar a sus sucesores. En otros términos: la estabilidad y eficiencia requieren normas permanentes que trascienden la duración de una sola generación; la democracia y la participación requieren que cada generación tenga plena capacidad de decisión en todas las cuestiones que conciernen a sus miembros; pero ¿no es el poder instantáneo también un poder sin objeto?

Una generación "sofisticada", que asume que todas las generaciones siguientes serán igualmente sofisticadas, adopta un enfoque del problema que es el de la teoría de los juegos, y busca una secuencia s_i, como se definió antes. Phelps y Pollak han demostrado[85] que en algunos casos esta secuencia de equilibrio existe e implica (en cierto sentido) subahorro y

Pareto-suboptimalidad. O sea, cada generación adoptará la misma proporción de ahorro s_1 mientras exista una relación s_2 ($s_2 > s_1$) tal, que todas las generaciones derivarían mayor utilidad si todas adoptaran s_2 que cuando todas adoptan s_1. Tendríamos que observar aquí que la relación s_1 no es una estrategia dominante, en contraste con el juego que resulta cuando la utilidad de cada generación está en función exclusivamente de su propio consumo pues, entonces, la relación cero de ahorro domina sobre todas las demás relaciones. En este caso, la realización de la solución en un juego sin estrategia dominante es más factible que en la mayoría de las demás relaciones; porque una vez que una generación adopta el punto de vista estratégico, puede razonablemente esperarse que sus sucesores también lo adopten. Phelps y Pollak hablan de comer la fruta de la sabiduría. La sofisticación es contagiosa en el tiempo, si no lo es en el espacio.

(d) En algunos casos, las suboptimalidades estática y dinámica pueden ser consideradas conjuntamente, como reforzándose o anulándose una a otra, o como alternativas mutuas. Primero, menciono brevemente un ejemplo en el que ambas trabajan en la misma dirección, por ejemplo, la inversión en una investigación fundamental. Para un país individual, esencialmente si es pequeño,[86] es preferible ser un país conductor y explotar los avances específicos hechos en otro lugar, por lo menos hasta donde pueda sin hacer más que una módica investigación. También, las inversiones hechas por una sola generación dependerán muy fuertemente de las preferencias temporales del gobierno. Por la falta de correlación en espacio y tiempo entre la inversión en investigación básica y lo que esta inversión produce, el chauvinismo y la miopía, o ambos a la vez, provocarán que exista subinversión y suboptimalidad.

(e) En otros casos, los dos mecanismos pueden trabajar en direcciones opuestas, como en el "imperialismo de libre comercio".[87] Supóngase que existen dos grupos de países, grupo A y grupo B, y que por algún motivo B no puede formar un cártel. El problema, entonces, es si los miembros de A tendrían que comerciar con los países de B sobre una base no cooperativa (competitiva) o sobre una base cooperativa (precios de cártel). En cualquier momento dado, las ganancias agregadas del grupo A serán maximizadas por la adopción de estrategias de cártel, incluso si para un país individual de A es mejor ser alguien que se corta solo y romper el cártel. Suponiendo que un cártel estable se mantiene de alguna manera esto puede verse como estrategia suboptimal en el tiempo, porque los países de B que enfrentan el monopolio y los precios monopólicos estarán motivados para desarrollar sus propias industrias y refinerías, si pueden hacerlo por precios inferiores a los mundiales, aun si al principio carecen de los recursos o la sofisticación tecnológica para lograrlo a precios competitivos. Esta es una variante del argumento de la industria incipiente, siendo la diferencia que los gobiernos y las industrias de B pueden ser tan miopes que no emprendan la producción si no les es *inmediatamente* factible producir a precios inferiores a los del mercado mundial.

Las suposiciones precedentes implican que para maximizar las ganancias agregadas a largo plazo los países del grupo A tendrían que adoptar la estrategia competitiva; en cualquier momento estarán tentados por salirse del cártel intertemporal debido a los precios bajos y formar un cártel instantáneo buscando precios altos, para maximizar a corto plazo las ganancias agregadas. Esto, sin embargo, puede ser anulado por la tentación de un país dado por romper el cártel instantáneo y utilizar los precios competitivos para maximizar las ganancias individuales a corto plazo. El comportamiento cooperativo en los juegos de largo plazo coincide con el comportamiento no cooperativo en los juegos de corto plazo, de manera que el nacionalismo y la miopía se anulan mutuamente. El "imperialismo del libre cambio" también demuestra en lenguaje sociológico que "el universalismo es el particularismo de los ricos". Me gustaría añadir que, incluso si la utilización de precios competitivos *puede* ser el resultado de un maquiavélico "cártel anticártel", puede también ser (y normalmente lo será) un resultado menos sofisticado de una situación atomizada de mercado. El lector recordará, en este punto, la falacia estructuralista antes criticada: del hecho de que las naciones se comporten de manera tal que maximizan sus ganancias a largo plazo no podemos concluir que esto es lo que las mueve a actuar.

(f) Podemos referirnos a los ejemplos que anteceden como "Dilemas Jerárquicos del Prisionero", en el sentido de que estamos enfrentando un juego intertemporal entre actores que están involucrados en subjuegos internos. La suboptimalidad estática y la dinámica pueden también interactuar de manera diferente, como *alternativas* mutuas, en el tipo de dilema básico de los arreglos institucionales del capitalismo. Basados en la visión de Joseph Schumpeter y Kenneth Arrow,[88] mencionaremos dos casos. Primero, el sistema de patentes, al internalizar los beneficios económicos de un invento, es dinámicamente óptimo en el sentido de maximizar la creación de nuevas posibilidades tecnológicas, pero al mismo tiempo es estáticamente subóptimo en el sentido de impedir la mejor explotación de las posibilidades existentes.[89] Segundo, un dilema similar entre la explotación y la creación de posibilidades surge de la aversión al riesgo del empresario. Maximizando la utilidad más que las ganancias, impide la realización del óptimo de Pareto estático, pero a la vez es su identificación con la empresa (y la consiguiente maximización de utilidad más que de ganancias) lo que hace del empresario la figura heroica del capitalismo desde el punto de vista dinámico.[90]

(g) En una prolongación directa de esta discusión, podemos intentar descifrar el sentido de la tercera y última de las principales apariciones del término "contradicción" en las obras de Marx: la "contradicción entre las fuerzas de producción y las relaciones de producción".[91] Por empezar, Marx nunca utiliza la frase favorita del marxismo vulgar: "la contradicción entre los medios sociales de producción y las relaciones privadas de producción" que, en sí misma, no es más contradictoria que un matrimonio entre una mujer alta y un hombre pequeño. La contradicción, si es una contradicción, no tendría que apoyarse en esta prestidigitación verbal. Segundo, en la

frase a interpretar, todos los términos —y no sólo el término contradicción— son muy difíciles de comprender. Simplificaré el problema estipulando que las fuerzas productivas sólo son tecnología, y las relaciones de producción, las distribuciones institucionales de la propiedad, dejando así de lado la tendencia de Marx a hablar como si todo (las relaciones de producción incluidas) fuera en cierto sentido una "fuerza de producción" si tiene un impacto en la eficiencia productiva.[92]

En el párrafo posiblemente más conocido (o casi) escrito por Marx, él afirma —dada la interpretación que acabamos de hacer de algunos términos clave— que en cada modo de producción las relaciones de propiedad son inicialmente "formas de desarrollo" de la tecnología, convirtiéndose, más adelante, en "frenos" de dicho desarrollo. La última expresión sugiere fuertemente algún tipo de suboptimalidad, sea de la explotación estática o del desarrollo dinámico de la tecnología. Desafortunadamente, es muy difícil comprender qué pensaba Marx, aun si nos limitamos al modo de producción capitalista. De acuerdo a lo que conozco, sólo un texto señala concretamente un mecanismo posible y allí Marx afirma que "En una sociedad comunista habría un campo muy diferente para el empleo de la maquinaria al que puede haber en la sociedad burguesa",[93] porque en la primera la maquinaria sería utilizada en todos los casos en los que el valor de trabajo de la máquina fuera menor que el valor creado por los trabajadores a los que reemplaza, mientras que en la última, la maquinaria solamente es empleada en los casos en que el valor del trabajo de las máquinas es menor que el valor de la fuerza de trabajo de los obreros reemplazados por la maquinaria. Esta suboptimalidad estática parecería ser, sin embargo, una característica permanente del capitalismo y no —como lo requiere el argumento— una característica que sólo aparece (o lo hace de manera creciente) en las últimas fases del capitalismo.

La interpretación más natural desde el punto de vista del economista moderno es el siguiente argumento: el capitalismo competitivo ofrece grandes incentivos para las invenciones; la nueva tecnología requiere tipicamente de producción en gran escala y utilización extensiva de la maquinaria; las economías de escala resultantes crean una tendencia hacia el oligopolio; el oligopolio, a su vez, reduce el incentivo de la invención. De estas cuatro proposiciones, la primera y la última están avaladas parcialmente por Arrow.[94] La tercera proposición es difícil de refutar ya que no es posible simultáneamente incrementar los réditos a escala y obtener maximización de ganancias y competencia perfecta. La segunda proposición es más dudosa, porque abstrae las innovaciones de la posibilidad de ahorro de capital. Por encima de este problema, sin embargo, hay una dificultad exegética, que Marx formula explícitamente:[95] bajo el capitalismo el progreso técnico se dará a un ritmo siempre creciente, más bien que decreciente.

Una última interpretación argumentaría que —como en el párrafo anterior— la suboptimalidad dinámica no debería entenderse como una declinación en la tasa absoluta de cambio de la tecnología sino como

relativa a la tasa de cambio que hubiera sido posible, en una misma etapa de desarrollo de las fuerzas productivas, en una organización comunista de la sociedad. Me parece que esto se acerca más a lo que Marx pensaba. Los detalles del argumento no están trabajados, pero aproximadamente podrían haberse elaborado así: en condiciones de bienestar material, la innovación es una actividad natural del hombre y surge de la fuente misma de su ser, pero en condiciones de escasez y pobreza, se requieren incentivos para que haya innovaciones. En los estadios tempranos del capitalismo, había inevitablemente gran escasez y pobreza ya que faltaban las condiciones materiales para satisfacer las necesidades universales. En esas condiciones, el sistema de patentes era mejor y más progresista, aunque subordinara el progreso a las ganancias. Este sistema, sin embargo, creaba inevitablemente las condiciones para su propia extinción. En fases posteriores del capitalismo hay todavía gran pobreza pero es evitable. Dada la tecnología desarrollada por el capitalismo mismo, es materialmente posible instalar un régimen en el que el nivel de bienestar material sea lo suficientemente alto como para que la innovación surja como una actividad espontánea, aun a un ritmo muy superior a todos los anteriores. En esta interpretación, por lo tanto, es posible caracterizar el ritmo de cambio técnico bajo el capitalismo como creciente y subóptimo.[96]

No me detendré en los méritos intrínsecos de este argumento, excepto para afirmar que, personalmente, tiendo a aceptarlo. Por otra parte, no acepto el corolario implícito de Marx, de que esta "contradicción entre las fuerzas y las relaciones de producción" puede también servir como palanca para el cambio social. Sostengo que si la brecha entre lo potencial y lo real actúa como una fuerza social en una sociedad dada, tiene que coincidir con la brecha entre lo real de esa sociedad y lo real en otra, todavía no realizado en la primera. Desde luego, no niego el poder del pensamiento Utópico en los movimientos sociales, pero creo que representaría mal a ese pensamiento verlo como una teoría de la suboptimalidad. Cuando el hombre lucha por realizar una Utopía no lo hace por realizar lo óptimo relativo a preferencias dadas sino, con mayor frecuencia, lo hace para cambiar al hombre y a sus preferencias. Este rasgo de pensamiento es también muy importante en el marxismo, pero para los fines de la presente discusión me concentro —como creo que Marx mismo lo hizo en la teoría ahora en discusión— en la ineficiencia del capitalismo y no en la perversión de sus carencias.

Para reformular el argumento por la negativa, no creo que una suboptimalidad definida como relativa a una mera posibilidad no realizada en ningún lugar pueda ser un agente importante en el cambio social. Alexander Gerschenkron trata bien el punto en un análisis comparativo de la industrialización europea, particularizando como una variable clave a "la tensión entre el estado real de las actividades económicas en el país y los obstáculos existentes para el desarrollo industrial, de un lado, y la gran promesa inherente a dicho desarrollo, de otro lado."[97] Esta tensión o suboptimalidad no entra en el análisis de la industrialización *inglesa*, sino sólo en las formas asumidas por la industrialización de Alemania, Italia, Rusia,

etcétera. Una vez ocurrida la Revolución Industrial en Inglaterra, sus resultados podrían aparecer como "promesas" para otros países y definir para ellos la brecha entre lo real y lo potencial, pero, para el primer país no se puede hablar con utilidad de dicha brecha. Una vez que el comunismo ha sido introducido con éxito en un país, la brecha entre los logros de ese país y la triste realidad de otros países podría servir como una palanca para la revolución social en éstos. Pero la *primera* y crucial revolución comunista no puede darse de esta manera. (Véase también la falacia de composición inherente al pensamiento difusionista.)

(h) Mi último ejemplo de suboptimalidad es el modelo capitalista como juego diferencial exhibiendo suboptimalidad,[98] de Kelvin Lancaster. En un juego diferencial de dos personas, cada actor controla, dentro de ciertos límites, el valor de una variable x_i (t), i = 1,2. Una estrategia es un perfil temporal x_i (t) de una variable definida por un período (t_0,t_1). El resultado puede definirse de diversas maneras, pero aquí tratamos el caso en el que el resultado es una función acumulativa de estrategias empleadas hasta el punto en el tiempo en que se llega al resultado. Presumimos que ambos actores desean maximizar el resultado del período total (t_0,t_1). Las nociones de un punto de equilibrio y una solución se definen análogamente a los del caso estático, para el cual véase el apéndice 1. Debemos tener cuidado, sin embargo, de interpretar a una estrategia como mejor que la otra: contra una estrategia dada x_2 (t), x_1 (t) es mejor que x'_1 (t), si da un mayor resultado al final del período, lo cual no implica un mayor resultado acumulado en cada momento durante el período.

En el modelo de Lancaster se supone que los trabajadores controlan la proporción entre el consumo obrero actual y la producción, ya sea directamente, por el control del nivel salarial o, indirectamente, controlando el impuesto a las ganancias. Ese control tiene límites por la altura y por la base, que arbitrariamente pueden estar cercanos al 100% y 0, respectivamente. Los capitalistas controlan qué proporción del producto no consumido por los trabajadores debe ser invertida y cuál la utilizada para el consumo capitalista, con restricciones similares. Para las economías capitalistas modernas este modelo me parece superior al tradicional, de Marx, porque une varias facetas de la explotación. En el capitalismo de la época de Marx, el límite superior de la variable del control de los trabajadores era muy bajo y, en sí mismo, esto parecería una forma de explotación. (Presumo que este límite superior está establecido por una relativa negociación de poder entre las dos clases.) La explotación es especialmente severa si los capitalistas eligen un valor bajo para la tasa de reinversión de sus ganancias, pero argumentaría, sin embargo, que debemos hablar de explotación incluso si una gran parte del sacrificio impuesto a los trabajadores es reinvertido para el beneficio de los trabajadores futuros. En el capitalismo moderno, el límite superior de la variable de control de los trabajadores se encuentra en un nivel más alto, pero como los trabajadores no controlan el proceso de inversión tienen que elegir un valor para el control muy por debajo de dicho límite; sacrificio elegido más bien que impuesto, o impuesto por la estructura del juego más que por la clase capitalista. El hecho crucial es que los

trabajadores tienen que mantener su variable de control muy por debajo del límite superior para que haya inversión y un incremento en el consumo futuro, incluso si no están seguros de que los capitalistas van a reinvertir el excedente más que a gastarlo en el consumo inmediato. Un dilema similar rige para los capitalistas, pues si no invierten parte de las ganancias no tienen ninguna garantía de obtener más tarde su participación en el producto futuro incrementado.[99]

Ahora, el análisis matemático de este juego (o más bien de un caso especialmente importante) demuestra que tiene una solución con diversas características interesantes. Primero, los perfiles temporales de las variables de control son discontinuos; existe un tiempo t ($t_0 < t < t_1$) tal que hasta t ambas clases consumen mínimamente, mientras que a partir de t, ambas clases consumen máximamente. Segundo, la solución es Pareto-subóptima: existen $x_1(t)$ y $x_2(t)$ que dan un consumo mayor para ambas clases que el consumo definido por las estrategias de solución, siendo el obstáculo que $x_1(t)$ y $x_2(t)$ no son óptimas una contra la otra. Lancaster ha demostrado que la característica de suboptimalidad se preserva incluso si introducimos diversas modificaciones para hacer el modelo más realista. No estoy totalmente seguro de que estas modificaciones sean suficientes para hacer del modelo más que una aproximación estilizada a algunos aspectos del capitalismo moderno; no veo nada en particular en la economía real que corresponda a la característica *bang-bang* de discontinuidad exhibida por la solución del juego. Sin embargo, el modelo sugiere la muy importante idea de que en las economías modernas de tipo mixtas, los trabajadores tienen el poder político y los capitalistas, el económico; los trabajadores el control sobre las decisiones de ahorro y los capitalistas, sobre las decisiones de inversión; *los trabajadores, el poder sobre el presente, y los capitalistas, sobre el futuro*. Sugiere, también, que en algunos casos cuya importancia es incierta, la separación de poderes es Pareto-subóptima. Desde mi punto de vista, sin embargo, la recriminación al capitalismo surge de la desigualdad de acceso a las decisiones de inversión y no de la ineficiencia, que puede ser o no, producto de esta desigualdad.

Una teoría dual del cambio social

En la última sección de este capítulo uniré algunas de las tendencias mostradas en las páginas precedentes en una teoría —o, más modestamente, en un esquema explicativo— del cambio social. El argumento versa sobre la relación entre cambio, contradicciones, condiciones estructurales y acción política. La tesis principal es que, dadas ciertas condiciones estructurales tales como la proximidad espacial de un grupo, con bajas tasas de cambios dentro del grupo,las contradicciones tienden a generar una acción colectiva con el fin de superar las contradicciones. Esta acción colectiva puede tener el efecto de revertir o detener el proceso de cambio, por ejemplo, si la contradicción a ser remontada es en sí misma un vehículo de cambio. Si, por el contrario, las contradicciones son configuraciones

estables de la sociedad, el proceso político podrá traer consigo un cambio, si es exitoso. Como la acción colectiva puede o bien prevenir o bien generar el cambio, las condiciones para dicha acción pueden funcionar tanto como condiciones para la estabilidad como para el cambio. En este sentido, propongo una teoría *dual* del cambio social.

No es difícil ver cómo algunas de las contradicciones a las cuales me he referido constituyen en sí mismas una forma de cambio. Es el caso de todos los ejemplos de contrafinalidad y, en verdad, creo que esta articulación vale prácticamente para todos los casos.[100] Si la erosión, una tasa descendiente de ganancia o una fuerte crisis de demanda no son ejemplos de cambio social, nada lo es. Podemos aquí hablar de *una variedad económica del cambio*, resultante de la interacción estadística de innumerables elecciones individuales. La mayor parte de los casos de suboptimalidad, por su parte, no son en sí mismos casos de cambio. En verdad, esto es correcto para la no provisión de bienes públicos o la ausencia de organizaciones. No es verdadero para las malas economías externas, tales como la polución o "la tragedia de las comunas", y no evidentemente verdadero ni falso para los demás ejemplos de suboptimalidad antes discutidos. Para simplificar, hablaré como si las suboptimalidades nunca constituyeran un cambio y como si la contrafinalidad siempre lo hiciera, lo cual implica, entre otras cosas, que un externalidad será clasificada como un ejemplo de contrafinalidad (lo cual, en verdad, frecuentemente es). Dada una suboptimalidad, entonces, una acción colectiva para obtener el óptimo de Pareto representará la *variedad política del cambio*, un alejamiento del *status quo*. Pero, dada una contrafinalidad, la finalidad de una acción política será restablecer el *status quo* y, por lo tanto, contrarrestar al cambio.

Este es el esqueleto del argumento, que después debe ser llenado de carne. Como en la *Fenomenología del espíritu* de Hegel, las respuestas a las contradicciones pueden ser ordenadas jerárquicamente, comprendidos algunos *niveles* y *modos* de aprendizaje. Hegel habla del itinerario del espíritu como una serie de *Momente* (conciencia, autoconciencia, razón, espíritu y religión) dentro de cada uno de los cuales pueden diferenciarse varias *Gestalten*. Dentro de cada *Moment* o modo de conciencia, cada *Gestalt* o nivel representa una forma particular de la contradicción específica de ese momento. Los deseos contradictorios del amo, por ejemplo, son sólo una forma particular de la contradicción general que define el modo de autoconciencia y que podría, en términos aproximados, caracterizarse como el deseo de tener la torta y comérsela. En cada intento por trascender la contradicción a un nivel dado, es creada una nueva contradicción hasta que, finalmente, se ve que el modo, como un todo, es contradictorio, visión que conduce a que surja un nuevo modo . (La noción del aprendizaje jerárquico e inherentemente contradictorio se encuentra, también, en el trabajo de Gregory Bateson sobre el "deutero aprendizaje", en el que sugiere que "la criatura es conducida a [aprender en] el nivel III por los 'contrarios' generados [por el aprendizaje] del nivel II".[101]

También hemos visto trabajar a esta lógica en el análisis del ciclo de la telaraña. En el nivel cero de aprendizaje, simplemente nos comportamos

de acuerdo con los programas de oferta y demanda; puede decirse que este "aprendizaje"es exitoso si las oscilaciones desaparecen lo suficientemente rápido. En el primer nivel de aprendizaje tenemos, entonces, el comportamiento generado por la presunción de que todos los otros productores actúan de acuerdo con sus programas de oferta y demanda; en el nivel *n* de comportamiento generado por la presunción de que todos los demás productores están en el nivel *n*-1. En cualquier nivel de la jerarquía las oscilaciones pueden deprimirse lo suficiente como para que el proceso llegue al final, pero también es posible que los actores se den cuenta de que la jerarquía misma es inherentemente contradictoria, y que todos se encuentran en el mismo nivel. En ese caso, tenemos una transición a un nuevo *modo*: el estratégico o el de la teoría de juegos. Sin embargo, si este modo también exhibe alguna variedad de contradicciones, sólo el paso al (tercero y último) modo de acción colectiva puede superar la contradicción. En la filosofía de Hegel, todas las contradicciones finalmente se reconcilian en el modo de la Sabiduría Absoluta; sugiero que el equivalente en el marxismo sería la acción colectiva organizada. Los tres modos son, entonces, definidos respectivamente, por una presunción compartida de ventaja unilateral, una presunción compartida de racionalidad estratégica y la solidaridad de la acción organizada. Como se señala en el apéndice 2 de este capítulo, la racionalidad estratégica es en un sentido un vehículo de la libertad colectiva, pero solamente desde un punto de vista formal. Desde un punto de vista sustancial, el Dilema del Prisionero nos dice que, incluso en la racionalidad transparente de la teoría de los juegos, los actores pueden estar forzados a actuar colectivamente de manera desastrosa. Existe la frustración de estar atado a fuerzas causales que uno no entiende, pero puede ser igualmente intolerable entender que uno está deliberadamente contribuyendo al estado Pareto-inferior. Para sintetizar, el primer modo de existencia social, necesariamente, genera contradicciones que pueden o no ser eliminadas por la transición al segundo modo; si no lo son, sólo el paso al tercer modo puede resolver el dilema. Es innecesario decir que esta secuencia es solamente lógica. En cualquier caso real de cambio, uno o más pasos pueden saltearse.

En este punto se requieren algunas palabras para explicar el *papel de los juegos sin solución no cooperativa*, a los que se hizo referencia en la sección anterior. Estos juegos son principalmente de dos tipos: juegos sin punto de equilibrio y juegos con varios puntos de equilibrio, ninguno de los cuales puede ser escogido como la solución. (Para estas ideas véase el apéndice 1 de este capítulo.) Como ejemplo de un juego sin un punto de equilibrio podemos quizá tomar la hiperinflación, donde el rédito mayor va al jugador que prevé y actúa sobre un mayor aumento de precios que el previsto por los demás. Un modelo formal de este juego es el siguiente: "Cada jugador escribe un número, el jugador que escribe el número mayor recibe de los otros una suma de dinero igual a la diferencia entre el número mayor y el escrito por ese jugador." Es intuitivamente claro y puede probarse rigurosamente,[102] que en esta *fuite en avant* puede no haber un punto de equilibrio y, *a fortiori*, ninguna solución. Probablemente, los jue-

gos con diversos puntos de equilibrio pero sin solución son más importantes. La provisión de bienes públicos puede servir como ejemplo. Si hay muchos actores de aproximadamente el mismo tamaño involucrados, la provisión de bienes públicos es una *n* persona del Dilema del Prisionero. Si entre los actores hay uno más grande que los demás, puede darse cuenta de que le es útil proveer el bien en cuestión, aunque no impedir que los demás se corten solos; a esto nos referimos con anterioridad como la explotación del grande por el pequeño. Finalmente, si hay dos actores grandes, cada uno de los cuales actuaría como proveedor único del bien si estuviera seguro de que nadie lo haría, tenemos un juego de la Gallina, es decir, un juego sin solución. Ya hemos observado que el problema del duopolio y el análisis de Marx-Fellner de la tendencia al cambio técnico conducen a juegos similares sin solución: los ejemplos más conocidos son negociar por medio de amenazas, la carrera armamentista, etcétera.

Los juegos sin solución son *sui generis*, comparten algunas características, pero sin coincidir, con la contrafinalidad y la suboptimalidad. Como no hay nada racional que hacer, los actores *tienen* que formular alguna hipótesis psicológica acerca de lo que los demás posiblemente harán, incluso si todos son conscientes de que tratan de adivinar lo que los otros piensan. Estas hipótesis, por supuesto, pueden ser contradictorias y generar contrafinalidad. Sostengo que ésta es una forma particularmente intolerable de contrafinalidad, porque cada actor sabe que los otros son racionales y, a pesar de ello, debe tratarlos como si no lo fueran. En el modo preestratégico, cada participante actúa sobre la creencia sincera (aunque ingenua) de que él está un paso adelante de los demás, pero la contrafinalidad posestratégica encontrada en los juegos sin solución es más perversa, porque todos son plenamente conscientes de la posibilidad de ser adivinados. Incluso, el comportamiento casual, por cierto, puede no ser importante porque la ausencia de una solución cubre el caso de estrategias combinadas. Por lo tanto, en los juegos sin solución la presión hacia un comportamiento cooperativo tendría que ser especialmente fuerte. "En los casos en que el concepto de equilibrio no conduce a un modo único de comportamiento, los jugadores probablemente hacen bien en tener en cuenta el juego cooperativo."[103] Siempre que más abajo hablo de contradicciones incluyo a los juegos sin solución no cooperativa que, bien puede decirse, constituyen una tercera variedad de contradicciones. No creo que sea comparable en importancia a las otras dos variedades. Durante mucho tiempo se pensó que la clase de juegos sin solución era mucho mayor que lo que realmente es,[104] pero la introducción por Von Neumann de la noción de estrategias combinadas aumentó ampliamente el número de casos en los que "el concepto del equilibrio no conduce a un modo único de comportamiento".

Ahora, afirmaré y discutiré las cuatro proposiciones que forman la sustancia de mi enfoque:

(1) *Si los actores perciben que se están comportando de manera contradictoria, intentarán organizarse con el fin de superar la contradicción.*

Me extenderé sobre esta idea, que en primera instancia puede parecer trivial (o vacía). Una noción similar fue propuesta por Arrow, quien sugiere que "cuando el mercado fracasa en lograr un estado óptimo, en cierta medida la sociedad reconocerá la brecha y surgirán las instituciones sociales no mercantiles para intentar tender un puente".[105] En *Exit, Voice and Loyalty*, Albert Hirschman aprueba y cita este pasaje, después del siguiente párrafo de su anterior trabajo sobre *The Strategy of Economic Development*:

> La tradición parece requerir que los economistas discutan permanentemente si en cualquier situación de desequilibrio *las fuerzas del mercado, actuando por sí mismas*, podrían restaurar el equilibrio. Pero, en verdad, como científicos sociales debemos plantearnos un problema más amplio: ¿puede corregirse la situación de desequilibrio por las fuerzas del mercado y por las que no son del mercado, o por ambas actuando conjuntamente? *Afirmamos que las fuerzas no mercantiles no son, necesariamente, menos automáticas que las del mercado.*[106]

Otro enfoque fuertemente influido por la observación de Arrow es la "teoría del cambio institucional" diseñada por Douglass North y Lance Davis, quienes "postulan que las instituciones económicas son renovadas o los derechos de propiedad son revisados porque parece beneficioso para los individuos o grupos sobrellevar los costos de esos cambios; esperan obtener algún beneficio o ganancia, imposible de lograr bajo el viejo acuerdo."[107] Estas referencias parecen probar que, de todas maneras, la discusión de la primera proposición anterior no es una simple pérdida temporales.

Podemos observar que todos estos autores se ocupan principalmente de la suboptimalidad, especialmente la carencia de organización y la no provisión de bienes públicos. Solamente Hirschman, al pasar, se refiere a la intervención política en el ciclo de los negocios. A pesar de la desorientadora palabra "automáticas" en el citado pasaje de Hirschman, estos autores son muy cuidadosos al señalar que los fracasos del mercado no *siempre* generan una acción política exitosa para obtener lo óptimo o el equilibrio. De hecho, North y Davis hacen más que formular un alerta general contra esta concepción errónea al intentar especificar las condiciones que tienden a impedir dicha acción. Vuelvo más adelante sobre algunas de sus ideas. Aquí quiero enfatizar, tanto como sea posible, que la primera proposición no implica una visión cibernética de la sociedad como un sistema ultraestable.[108] El pasaje de Hirschman, si se lee sin el alerta que le sigue, parece implicar que el proceso político es un proceso regulatorio de segundo orden, que se pone en marcha cuando el mecanismo de primer orden del mercado fracasa en alcanzar lo óptimo o el equilibrio. Este punto de vista es sostenido por la ahora extinta escuela de teóricos sistémicos en

ciencias políticas,[109] pero hoy no es tomado muy en serio o, por lo menos, así lo creo. Hay, en verdad, algunos automatismos incorporados a las sociedades, tales como las casi automáticas medidas que aparecen cuando la inflación o las tasas de desempleo exceden ciertos niveles, pero el proceso que estableció esos mecanismos estaba, en sí mismo, lejos de ser automático. Uno se pierde la esencia de la historia si cree que la creación de nuevas instituciones comparte el carácter "automático"del funcionamiento cotidiano de esas instituciones una vez establecidas. (Véase también la diferenciación entre política rutinaria y política creativa, discutida en el capítulo 3.) Lejos de ser automático, el intento por alcanzar Pareto-optimalidad o para restablecer el equilibrio puede fracasar totalmente, por razones a ser discutidas más adelante.

Por supuesto, es crucial en este contexto que las contradicciones sean *percibidas*. Si los granjeros (en el ciclo de la telaraña) culpan al clima por las fluctuaciones cíclicas de los precios, no van a hacer mucho, excepto hablar acerca de ello. También enfaticé antes que la ausencia de un bien público tiene un poder motivante y movilizador sólo si es una ausencia *sentida* y que típicamente significa la ausencia de algo que está presente en otro lugar. También he mencionado al pasar que para que el problema del que se corta solo en la formación de sindicatos sea un *problema*, los trabajadores deben percibir que la adopción conjunta de la estrategia cooperativa es Pareto-superior a la carencia de organización y que el juego del salario no es de suma cero. No discutiré por separado las condiciones para que las contradicciones sean percibidas, ya que creo que en gran medida coinciden con las condiciones para la organización, que discutiré más adelante.[110]

El problema de la suboptimalidad dinámica debe tratarse por separado, ya que la noción de generaciones sucesivas que se unen para un ataque organizado contra el nivel de inversión subóptimo, es poco creíble, por decir lo mínimo. Aquí, la transición al comportamiento cooperativo tiene que ser llevada a cabo por un solo actor (por ejemplo, una sola generación) con la esperanza de que los sucesores le den continuidad. Los individuos que pertenecen a la generación crucial no ganarán nada con el paso a la cooperación, mientras que en el caso estático todos experimentarán ganancia. No estoy muy seguro acerca de este problema, pero sugiero que dos conjuntos de condiciones contribuyen a la probabilidad del cambio. En primer lugar, la generación crucial tiene que satisfacer las condiciones descritas en las siguientes proposiciones 2 y 3. En segundo lugar, tiene que estar en una posición que obligue a sus sucesores a seguirla, por ejemplo, estableciendo una constitución que no pueda ser fácilmente (o legítimamente) derogada por las generaciones posteriores, para su beneficio. Si el bienestar de las futuras generaciones entra en la función de utilidad de la generación presente, entonces el sacrificio en aras de *todas* las generaciones posteriores será presumiblemente más atractivo que un sacrificio sólo por la siguiente generación. Sea como fuere, creo que la primera proposición concerniente a la *tendencia* a cooperar es válida también en el caso intergeneracional, aun si las condiciones para que esta tendencia se realice son más arduas.

Podemos también observar aquí que no es siempre sencillo discernir si estamos tratando con un caso de contrafinalidad (o de suboptimalidad estática) o con un caso de suboptimalidad dinámica. Tomemos la erosión: si los efectos de la deforestación no son sentidos en la actual generación, estamos tratando con un Dilema del Prisionero intertemporal, en el que el interés colectivo e individual de corto plazo requieren un comportamiento no cooperativo, opuesto al comportamiento cooperativo requerido por el interés colectivo de largo plazo. Sin embargo, si la duración del mecanismo causal es menor que el horizonte temporales de los individuos, los intereses colectivos e individuales van a divergir, aun a corto plazo. El mismo punto es tratado en el análisis de Marx sobre las Leyes de las Fábricas Inglesas, las que explicó en parte como el resultado de la actividad de los trabajadores y en parte como defensa de la clase capitalista contra sus propios miembros,[111] siendo la idea que la sobreexplotación de los trabajadores podría amenazar su reproducción física y, por lo tanto, al mismo capitalismo. La última explicación requiere que los intereses colectivos de los capitalistas superen sus intereses individuales o que sus intereses de largo plazo superen a sus intereses de corto plazo. Ambas ideas, sin embargo, son difíciles de encuadrar en lo que Marx dice, en algún otro lugar, sobre la imposibilidad de la solidaridad entre los capitalistas.[112]

Como última observación acerca de la posibilidad de sostener esta proposición sobre la suboptimalidad dinámica, podemos notar que en casos como el del imperialismo de libre cambio, la proposición indica que hay una tendencia tanto hacia la formación de un cártel como de un cártel-anticártel. Estas afirmaciones pueden ser verdaderas simultáneamente, incluso si no pueden realizarse simultáneamente (como, sin embargo, tendrían que hacerlo de acuerdo con la teoría cibernética). Aquí, las condiciones para una solidaridad intergeneracional no son tan arduas, ya que pueden aliarse con los intereses individuales a corto plazo.

(2) *La probabilidad del éxito en el intento de organizar varía inversamente a la distancia comunicativa entre los miembros.*

Pienso en la "distancia comunicativa" como función de distancia espacial y de tecnología del transporte, de manera que los miembros de una sociedad densamente poblada con comunicaciones escasamente desarrolladas pueden estar más aislados, unos de los otros, que los miembros de un país escasamente poblado con buen transporte. El fundamento detrás de la proposición (2) (y la proposición (3)) es la simple observación de que es necesaria la confianza para superar las contradicciones y que las personas tienen que estar juntas por algún tiempo para que lleguen a confiar unas en otras. Frecuentemente se ha mencionado la importancia de la proximidad espacial para la solidaridad del grupo. De acuerdo con los Escritos Federalistas, "El límite natural de una democracia es la distancia del punto central, que permita a los ciudadanos más alejados reunirse tan frecuentemente como lo exijan sus funciones públicas." En el *XVIII Bru-*

mario, Marx explica la falta de conciencia de clase entre los campesinos franceses por su existencia aislada. El surgimiento de la conciencia de clase —la transición de la clase en sí a la clase para sí— es, en verdad, un caso especial del principal problema aquí discutido, la transición de la maximización individual a la del grupo, o del comportamiento no cooperativo al cooperativo. En contraste con el aislamiento del campesinado, la proximidad física de los trabajadores en la fábrica aumenta su solidaridad; están "disciplinados, unidos, organizados por el mecanismo mismo de producción capitalista".[113] Marx también afirma que "La dispersión de los trabajadores rurales en grandes áreas quiebra su resistencia, mientras que la concentración aumenta la de los operarios urbanos",[114] y en términos muy similares a los de la proposición (2) también postula que "el poder de resistencia de los trabajadores disminuye con su dispersión".[115]

Se dirá —y yo concuerdo— que estas son nociones, en verdad, obvias. Es posible, sin embargo, avanzar un paso más observando que la organización de clase, además del objetivo de superar las contradicciones de la clase misma, tiene, también, el objetivo de perpetuar las contradicciones de la clase opuesta. Esto puede hacerse, por cierto, dejando fuera de la ley a las asociaciones de la clase trabajadora, como lo intentaron las clases gobernantes inglesas por medio de las *Combination Acts.* En general, se acuerda que estas Leyes tuvieron una dudosa eficiencia e, incluso, pudieron haber ayudado a desarrollar a las asociaciones que intentaban prohibir.[116] En otras situaciones, prohibir a los sindicatos o matar a los dirigentes de la clase trabajadora puede ser más eficaz, pero un expediente más simple y a menudo más inteligente es reprimir a la clase explotada controlando las condiciones estructurales pertinentes para el éxito de su organización. La resistencia a los ferrocarriles por parte de las clases gobernantes en el Sur esclavista[117] o en Rusia antes de 1860[118] puede explicarse en este sentido, como opuesto a lo manifestado implícitamente en un reciente artículo de investigación, en el que los sistemas de transporte son considerados como "agentes que facilitan la integración y el mantenimiento de la sociedad".[119] Si miramos la estructura de clase más bien que simplemente a la "sociedad", un sistema de transporte con efecto integrador para *todas* las clases puede tener un efecto negativo neto para las clases gobernantes (siendo la utilidad marginal del transporte mayor para las clases que antes tenían menor acceso al mismo) y, por lo tanto, un efecto desestabilizador sobre la sociedad en su conjunto. En el micronivel, el diseño espacial de las fábricas a menudo está concebido para minimizar las posibilidades de interacción prolongada entre los trabajadores. Usualmente esto se justifica en nombre de la eficiencia productiva, pero ciertamente tiene, también, el efecto, intencionado o no, de obstaculizar el desarrollo de la conciencia y la combatividad de la clase trabajadora.

Dos aparentes excepciones a la proposición (2) permiten avanzar en un desarrollo conceptual. En su importante trabajo sobre *The Yenan Way in Revolutionary China,* Mark Selden intenta refutar la hipótesis que re-

laciona la revolución con la urbanización (y, por lo tanto, con un sistema altamente desarrollado de transporte) observando que la Revolución China fracasó totalmente en las áreas más desarrolladas y fue, finalmente, exitosa en la remota provincia de Shensi, con escasas comunicaciones internas.[120] Selden admite, sin embargo, que el movimiento revolucionario *surgió* inicialmente en las áreas más desarrolladas y esta concesión puede ser suficiente para nuestro objetivo. Un ejemplo en contrario de la proposición (2) sería la demostración de que en países con comunicaciones uniformemente escasas son más proclives a las revoluciones que los países con, por lo menos, algunas regiones desarrolladas, pero no parece ser éste el caso. La utilización de un *promedio* de distancia comunicativa puede ser muy tosco para establecer una generalización válida, del mismo modo que el promedio de duración de pertenencia a un grupo no es una variable pertinente en la proposición (3). Un segundo problema es aportado por el análisis de Georg Simmel acerca de la distribución espacial de los judíos durante la Diáspora.[121] Al dispersarse por todo el mundo, el pueblo judío aseguró la supervivencia del grupo porque ninguna persecución única podía golpearlos a todos. Al mismo tiempo, los judíos en una ciudad dada se reunieron en guetos para poder defenderse. Este último hecho es acorde con mi argumento, pero el primero parece un contraejemplo, ya que la dispersión de los grupos minoritarios perseguidos puede ser una condición para la supervivencia y, por lo tanto, para la organización. La excepción, sin embargo, puede ser sólo aparente porque, en primer lugar, podríamos aplicar el mismo razonamiento que en la discusión del primer ejemplo en contrario y argumentar que la *decisión de dispersarse* no podría haber surgido en una comunidad dispersa. En segundo lugar, es posible que sea, precisamente, la verdad de la proposición (2) la que hace más probable la persecución en caso de concentración; pareciendo ser débil se puede preservar la poca fuerza que se tiene. (Esta línea de argumento contrario al contraejemplo fracasaría, sin embargo, si la concentración de las víctimas hace a la persecución más eficaz y no sólo más probable.)

La proposición (2) implica que un grupo pequeño tenga —*ceteris paribus*— más posibilidad de lograr lo óptimo que un grupo más grande. Algunas otras cosas, sin embargo, posiblemente no son iguales, porque la brecha entre el equilibrio no cooperativo y el cooperativo es más pequeña para grupos más pequeños.[122] En la medida en que la probabilidad de éxito se relaciona con la fuerza de la motivación y esta última con la ganancia potencial de la organización, los grupos pequeños tienen *menos* posibilidades de ser exitosos. No es fácil decir *a priori* cuál será el efecto neto de estas tendencias opuestas.

(3) *La probabilidad de éxito en el intento de organización varía inversamente al ritmo de cambio de la pertenencia de los individuos a un grupo.*

De nuevo, la proposición se apoya en la simple idea de que si los miembros del grupo que genera contradicción nunca permanecen en él un tiempo suficiente como para conocerse y confiar en los demás, es bastante difícil esperar que se sobrepongan a la contradicción. Creo que para las sociedades modernas éste es el obstáculo principal para el desarrollo de una organización exitosa, aunque el aislamiento espacial lo fue en el pasado. Las rebeliones estudiantiles que fracasan son hoy, la analogía de las rebeliones campesinas fracasadas ayer. También creo que las mismas causas que han reducido la importancia del aislamiento espacial han aumentado el ritmo de cambio de pertenencia al grupo, de manera que la capacidad neta para la acción colectiva puede, o no, haber mejorado. Uno puede, en verdad, estar de acuerdo con Mark Granovetter cuando destaca la "fortaleza de los vínculos débiles"[123] y argumenta que la acción colectiva se generará más fácilmente en las comunidades donde la mayor parte de las personas tengan muchos amigos que no se conocen entre ellos (y entre los que sirven como puente) que en las comunidades que puedan dividirse en agrupaciones aisladas de personas que se conocen muy bien entre sí. Esta correlación, sin embargo, está presumiblemente sujeta a una cláusula ceteris paribus que no siempre puede darse: que el número de vínculos fuertes tendría que ser el mismo en ambos casos. La movilidad se da entre las fuerzas que generan vínculos débiles pero también entre las que disuelven los fuertes, de manera que el efecto neto de la movilidad puede ser de ida y vuelta, incluso en el más favorable de los casos, tal como la movilidad geográfica en un área pequeña.

La idea que está detrás de la proposición (3) es frecuente e informalmente mencionada en la literatura, pero difícilmente se la discute en forma sistemática.[124] Un relevamiento del campo señala la necesidad de establecer por lo menos tres distinciones. En primer lugar, debe distinguirse claramente entre las condiciones para la *pertenencia* (digamos) al sindicato y las condiciones para la *formación* del sindicato. Mancur Olson observa que "La pertenencia compulsiva no puede...explicar la creación de los primeros, pequeños, sindicatos locales, pero sí puede responder por la viabilidad de los posteriores sindicatos nacionales, más grandes".[125] Seymour Lipset y Joan Gordon encontraron que había una correlación positiva entre la movilidad y la pertenencia al sindicato,[126] pero esto evidentemente no significa que los sindicatos podrían haber sido organizados, en primer lugar, por trabajadores de alta movilidad. La relación entre los estudios de la pertenencia a los sindicatos y los estudios de su formación es similar a la relación entre los análisis del equilibrio parcial y general, en la economía. El análisis de pertenencia tiene que tomar a los sindicatos ya dados y explicar cómo los trabajadores individualmente se relacionan con ellos, así como el análisis del equilibrio parcial toma los precios ya dados y examina cómo las firmas y los consumidores se adaptan a ellos. En la economía se reconoce desde hace mucho tiempo la necesidad de un análisis más general que también explique la formación de precios. De manera similar, yo sostendría que debería prestarse más atención al problema de la formación de los sindicatos.

En segundo lugar, el impacto de la movilidad sobre la acción colectiva puede ser estudiado en distintos niveles. Las correlaciones pueden encontrarse en los niveles del *individuo*, el *rol* o el *grupo*. Este último constituye aquí nuestra primera preocupación, pero son pertinentes unas palabras sobre los otros niveles. El muy conocido descubrimiento de que los individuos con tendencia a descender son mucho más conservadores que los del nuevo grupo que llega[127] está también formulado en el nivel del individuo. La hipótesis, formulada en el nivel de grupo, sería que el grupo que tenga una proporción mayor que el promedio de individuos con tendencia descendente sería también más conservador que el promedio, lo cual no implicaría ni surgiría de lo encontrado en el nivel individual. Un ejemplo de correlaciones en el nivel del rol es el análisis de Georg Simmel de los funcionarios elegidos en los Estados Unidos. De acuerdo con Simmel, esta institución presenta *prima facie* la paradoja de haber sido diseñada para servir a las personas, mientras que sirve realmente a quienes ocupan el cargo. Cuando los jueces, etcétera, son elegidos, tienden a considerar los cargos como una propiedad privada que pueden explotar para sus propios fines. Simmel sostiene que la paradoja desaparece cuando consideramos la alternativa a la elección de funcionarios: los funcionarios permanentes. Desde el punto de vista de las personas es mejor tener una clase gobernante inestable, que nunca puede consolidar su hegemonía, que una burocracia que adquiere un rol estructural, profesional y estable, incluso si el mal sistema inherente al primero no tiene analogía en el segundo. "El demócrata consecuente no será gobernado, incluso si esto significa que no puede tampoco ser servido."[128]

En tercer lugar, debemos distinguir entre "la simple experiencia de movilidad"[129] y la dirección de esa movilidad; entre los aspectos escalares y vectoriales, por así decirlo. El hecho mismo de estar en movimiento, social, geográfica u ocupacionalmente, puede tener un impacto independiente de si la movilidad es horizontal o vertical, o si vertical hacia arriba o hacia abajo. En la proposición (3) me preocupo principalmente del aspecto escalar de la movilidad. Veremos que, por lo menos en un caso importante, la cantidad y la dirección de la movilidad tienden a reforzarse mutuamente, pero no necesariamente es así.

El ritmo de cambio de la pertenencia, como variable del análisis social, es importante para establecer el alcance de los tiempos en muy distintos órdenes. Si es correcto, como sugirió, por ejemplo, Arrow, que las normas de comportamiento social, tales como los tabúes o hábitos,"son reacciones de la sociedad para compensar los fracasos del mercado",[130] sería de esperar que pasaran varias generaciones para que estas normas se establecieran con firmeza. El argumento puede generalizarse para cubrir las costumbres específicas de clase, tales como la solidaridad, y aquí la duración temporal se convierte en una variable fundamental. Por ejemplo, si una clase dada es una "clase de transición" cuyos miembros tienen típicamente padres e hijos en clases diferentes a la suya, entonces no adquirirá tradiciones y normas acumulativas. Un caso puro de dicha clase

de transición es el grupo de hombres liberados bajo la ley romana, "un estado evanescente restringido por ley a una sola generación".[131] (Este ejemplo también apunta al hecho de que una clase en transición bien puede ser una característica permanente de una estructura social si hay un ingreso constante a la clase en cuestión.) Un caso más importante es el de la clase media en el capitalismo moderno. Una famosa controversia entre Rosa Luxemburgo y Eduard Bernstein puede interpretarse en términos de la proposición (3). Bernstein demostró[132] que el número absoluto, e incluso el relativo, de pequeñas empresas capitalistas no tenía tendencia a declinar, contrariamente a lo que él tomó como la opinión de Marx. Luxemburgo replicó[133] que, en verdad, podía observarse la polarización predicha por Marx, pero en un sentido diferente al que Bernstein le adjudicaba. Incluso, si el número de pequeñas empresas era constante, el tiempo de vida de cada empresa individual sería cada vez más corto lo cual implicaba un metabolismo más rápido para toda la clase de los pequeños productores. Las nuevas posibilidades para las empresas pequeñas fueron creadas por el progreso técnico a un ritmo cada vez más rápido, pero las pequeñas empresas que surgieron fueron también absorbidas por grandes empresas capitalistas a un ritmo igualmente rápido, siendo la resultante una constante numérica, pero "evanescente" de la clase de los pequeños productores.

Por lo tanto, Marx estaba en lo correcto al predecir la disminución de la importancia política de la clase media: no por una declinación absoluta o relativa en su fuerza numérica sino por su carácter cada vez más pronunciado de clase en transición con escaso poder de organización autónoma. Esta es una analogía temporal con los obstáculos espaciales para el desarrollo de la conciencia de clase entre el campesinado francés, y no es sorprendente que marxistas posteriores[134] hayan utilizado al *XVIII Brumario* como un paradigma de la explicación del surgimiento del fascismo, con la clase media en el lugar de los campesinos y altas tasas de abandono de la pertenencia en lugar de la distancia geográfica. En estas interpretaciones, el bonapartismo y el fascismo sirven a los *intereses* del capital, aunque basándose en el apoyo popular del campesinado y la clase media, respectivamente. También, ambos regímenes han sido vistos como expresiones de la necesidad funcional del capitalismo de tener gobiernos aparentemente anticapitalistas. Como se mencionó antes, en esta explicación resulta difícil ver cómo la "necesidad funcional" se transforma en acción y en actores.

Muchos ejemplos pueden aducirse para demostrar el impacto del ritmo de cambio de pertenencia sobre la acción colectiva en tiempos más cortos. En el artículo antes citado sobre los levantamientos preindustriales, E.P.Thompson argumenta que el *mercado* ocupaba un lugar privilegiado en la génesis de las rebeliones."El mercado era el lugar donde las personas, que eran numerosas, se sentían por un momento fuertes."[135] *Por un momento*: el mercado no es un lugar donde la interacción prolongada pueda generar la conciencia de clase necesaria para una acción exitosa. Thompson demuestra cómo las revueltas por el pan casi nunca lograron sus objetivos inmediatos. Ya hemos cuestionado su implícita afirmación de que

las revueltas se perpetuaron por sus beneficios a largo plazo, lo que no niega que dichos beneficios puedan haber existido. Como Thompson observó, la regularidad de las revueltas por el pan puede haber actuado como impedimento y haber mantenido la explotación en un nivel más bajo que si no hubiera habido levantamientos. Tenemos aquí una interesante excepción de la proposición (3). En algunos casos, el valor bloqueador de un fracaso puede ser tan grande que confiere los mismos beneficios que los logrados por una organización plenamente lograda. Obviamente, el punto va contra una ambigüedad de la noción de intentos "exitosos" por organizarse. Si por éxito queremos decir que los objetivos del movimiento se logran, las revueltas por el pan son una excepción de la proposición (3), pero si también exigimos que los objetivos se logren *de manera intencionada*, el ejemplo en contrario desaparece. En el análisis de poder y posibilidad política del capítulo 3, se vio como noción clave *la producción intencional e inteligente de los resultados deseados*, y si este enfoque se adopta aquí, las revueltas por el pan y otros fenómenos similares (tales como el "éxito" de los movimientos estudiantiles) no constituyen una excepción a la proposición (3).

Otro contraejemplo aparente, el del "radicalismo" y la conciencia de clase, va contra una ambigüedad de otra noción. Se ha argumentado que, por lo menos en algunos casos, los segmentos más radicales de la clase trabajadora son los trabajadores itinerantes de la industria de la construcción,[136] lo cual pareciera ser incompatible con la proposición (3). Quisiera mencionar algunas nociones leninistas y argumentar que la"enfermedad infantil" del movimiento obrero es su ineptitud para dar "un paso atrás, dos pasos adelante". El radicalismo se define por el deseo impaciente de explotar al máximo las posibilidades del momento actual, a expensas de la creación óptima de posibilidades nuevas. Un movimiento obrero con plena conciencia de clase podría marcar el paso y hacer retiradas tácticas, mientras que los segmentos radicales urgirán hacia la acción inmediata.[137] Sin duda, el modo de existencia de los trabajadores itinerantes favorece la solidaridad de grupo y su capacidad para la acción a corto plazo, como las huelgas, pero esto no es lo mismo que la capacidad de *organizarse*. Para ponerlo en términos un poco diferentes: la solidaridad entre todos los trabajadores en un momento dado no es suficiente para la adquisición de ganancias durables; es necesario, además, que en momentos decisivos se desarrolle un sentimiento de solidaridad entre los trabajadores. El lector recordará la importancia de la solidaridad intemporal en el anterior análisis de la explotación: los dirigentes radicales de la clase trabajadora no podrán percibir las raíces históricas de la explotación, ni podrán entender por qué el "capital" y el "trabajo" no tendrían que estar a la par en cuanto a la remuneración.

Por consiguiente, no es verdadero que un modo de existencia itinerante genere siempre una actitud radical; sólo las condiciones especiales de los trabajadores de la construcción favorecen este desarrollo. En el caso típico, ir de un lugar a otro (geográficamente hablando) impide cualquier

desarrollo de la conciencia de clase, así sea para las ganancias (beneficios) a largo o a corto plazo. En los países en desarrollo, donde los trabajadores fabriles son campesinos que permanecen en la ciudad durante algún tiempo antes de regresar al campo, la formación de sindicatos se retrasa mucho. Una similar "hipótesis del ciclo vital" se ha formulado para la Noruega del siglo XIX y ligado al problema de la conciencia de clase.[138] Stephan Thernstrom ha discutido más o menos extensamente los "transeúntes permanentes" de la sociedad norteamericana "incapaces de enraizarse y de formar organizaciones".[139] Parece, sin embargo, que la discusión de Thernstrom se da en el nivel individual, más que de grupo. El no se ocupa, como lo hago yo aquí, de la identificación de grupos con una alta proporción de miembros con una corta duración de pertenencia, sino más bien de los individuos que casi nunca permanecen más de unos meses en un trabajo determinado. Por ese motivo es difícil comprender por qué la capacidad de "formar organizaciones" es relevante, a no ser que quiera decir que la proporción de estos transeúntes en toda la comunidad de la clase trabajadora era tan grande que *cualquier* tipo de sindicalización estaba retrasado.

La relativa infrecuencia o falta de éxito organizativo entre algunos grupos de las sociedades modernas encuentra una explicación, por lo menos parcial, en los altos ritmos de cambio de pertenencia que las caracterizan. Los reclutas, los estudiantes, los desempleados temporales,[140] los que buscan vivienda, los consumidores, los que utilizan el transporte público, los trabajadores inmigrantes, los turistas y los jubilados, todos entran en esta categoría de grupos constantes (o en crecimiento) numéricamente y pertenencia evanescente. En las palabras de Thernstrom, son "desorganizados, pulverizados, alienados, invisibles e impotentes".[141] Se ven como grupos sólidos en el modelo estadístico de cambio social de "montón de instantáneas", pero se evaporan rápidamente en el modelo longitudinal de "manojo de historias de vida". Tomemos como ejemplo típico las revueltas estudiantiles. Son verdaderamente frecuentes, pero casi nunca alcanzan la fuerza necesaria requerida para obtener resultados durables; no aprenden de sus errores anteriores ni aceptan la necesidad de marcar el tiempo, en algunos casos. Pienso que no es cuestión de que cada estudiante tenga un horizonte temporal que termina con su examen final. La suboptimalidad dinámica que resulta de la miopía política es una causa para organizarse, y no un obstáculo. El problema es más bien que los estudiantes no tienen suficiente tiempo para adquirir experiencia y desarrollar relaciones solidarias. Tres espacios temporales están comprendidos en dichos casos: la duración de la pertenencia individual al grupo, el tiempo mínimo requerido para el surgimiento de la solidaridad y el tiempo requerido para que los efectos negativos del comportamiento no cooperativo sean sentidos. Si el tercero excede al primero, tenemos una dinámica análoga a la del Dilema del Prisionero, pero si el segundo excede al primero, la dificultad es más bien la de cómo sobreponerse al dilema estático.

Haremos dos observaciones adicionales acerca del problema del *tiempo*. Primero, si los individuos del grupo tienen fuerte preferencia por el presente, su horizonte temporal efectivo puede ser aun más corto que la duración de pertenencia al grupo, de manera que esta última no es un obstáculo operante. En este caso, el alcance de la suboptimalidad dinámica es, por consiguiente, mayor. Segundo, una versión testeable de la proposición (3) tendría que especificar, entre otras cosas, cómo medir la duración de la pertenencia al grupo. ¿Tendríamos que tomar el promedio, la media o alguna medida como la proporción de una alta o baja pertenencia de los miembros del grupo? La elección, digamos, entre la duración promedio y la proporción de miembros del grupo de larga pertenencia comprende temas sustancialmente importantes. La existencia de una sólida minoría de miembros permanentes puede hacer más por la solidaridad que una duración promedio mayor con una pequeña dispersión alrededor de la media. El lector recordará un punto similar discutido en la proposición (2), el de que la existencia de algunas regiones con comunicaciones altamente desarrolladas puede tener más importancia para el potencial revolucionario que la existencia de un sistema mejor desarrollado y más uniforme.

En las páginas precedentes he intentado relacionar la acción colectiva con la "pura experiencia de movilidad" no importa en qué dirección. Como ya se ha mencionado, la dirección puede tener impacto adicional propio, que coincida o no con el impacto de la cantidad de movilidad. La interpretación tradicional, que en muchos casos puede ser correcta, ha supuesto que la movilidad hacia arriba de la clase obrera (para limitarnos a este importante caso especial) ha tenido efectos negativos sobre la solidaridad de la clase trabajadora, reforzando, por lo tanto, el efecto de la movilidad como tal. Dos mecanismos diferentes se han invocado para justificar dicha presunción. En primer lugar, que los organizadores potenciales de una clase son los primeros en abandonarla. "Mientras mayor sea la capacidad de una clase dominante de asimilar las mentes de vanguardia de una clase dominada, más estable y peligrosa se vuelve su dominación."[142] En segundo lugar —y ésta es la hipótesis más mencionada— la posibilidad de movilidad hacia arriba puede actuar como una válvula de seguridad, cuya eficiencia puede ser relativamente independiente del ritmo de movilidad real, como se observó en el capítulo 2. Aquí son relevantes los hallazgos de *El soldado norteamericano* explicado en el ejemplo de Boudon, puesto que si las aspiraciones de movilidad crecen más rápidamente que las posibilidades reales de movilidad, la movilidad hacia arriba puede actuar para incrementar el descontento, más bien que para reducirlo. No podemos entrar aquí en la discusión de este problema sino sólo observar que si el ritmo de movilidad es *muy* alto, puede presumirse que tanto el efecto neto como el efecto de cada mecanismo será el de reducir la capacidad de acción colectiva.

En esta conexión, la posibilidad de *contramovilidad*, para usar el término de Roger Girod,[143] es especialmente importante. Si, por ejemplo, los niños de la clase alta cuyos padres tuvieron padres de clase media se vuelven más frecuentemente de clase media que los niños de clase alta

cuyos padres vienen de hogares de clase alta, y así de manera similar para otras transiciones, cada movimiento de alejamiento de la propia clase tendería a ser opuesto al movimiento en la generación siguiente. Si esto se da en gran escala, puede haber una gran cantidad de movilidad aparente —suficiente para actuar como válvula de seguridad— pero movilidad de corta duración. En uno de los raros estudios de movilidad social durante tres generaciones, Ramkrishna Mukherjee concluyó que éste realmente es el caso. Su análisis sugiere que "la sociedad contiene un núcleo sólido de 'elementos estables' y, al mismo tiempo, un grupo de elementos 'inestables'. Estos últimos pueden comprender quizá sólo una pequeña proporción de miembros de la sociedad, pero son tan inestables que pueden verse envueltos en cambios de retroceso en su estatus, en generaciones sucesivas".[144] En esta afirmación parecería que algunas líneas familiares tendrían que ser inherentemente inestables. Una explicación alternativa de la información podría darse a través del mecanismo descrito por Marx en el texto arriba citado. Si pasar o permanecer en una clase diferente a la "propia" (una noción que tendría que precisarse) depende de las desviaciones de la capacidad promedio y si la capacidad es menos que perfectamente heredable, de manera que haya una regresión hacia el promedio, el resultado sería la contramovilidad.

Continuando ahora con algunas conclusiones sobre la proposición (3), podríamos observar que los procesos generales de incremento en el factor movilidad y en el cambio tecnológico han acercado más personas durante períodos cortos temporales, reduciendo la relevancia de la proposición (2) y mejorando la de la proposición (3). En las sociedades tradicionales, cada persona tenía acceso comunicativo permanente a un subconjunto muy pequeño de personas, cuyo comportamiento, juntamente con el suyo, generaba las contradicciones en las cuales él estaba atrapado. En las sociedades modernas, cada persona tiene acceso a todas las personas afectadas, pero el tiempo durante el cual este acceso es pertinente está restringido por la corta —y cada vez más corta— duración de la pertenencia al grupo.

Pienso que este marco conceptual permite dar una teoría de la "naturaleza objetivamente capitalista del Estado moderno". En la mayoría de los países capitalistas —y, en verdad, en la mayoría de los más avanzados— no se puede poner en duda la alianza subjetiva entre el Estado y los intereses del capital. Pero demasiadas actividades estatales están directamente opuestas a los intereses capitalistas y directamente adaptadas a los intereses de la clase obrera, para que una explicación conspirativa se mantenga a flote. Tampoco debería darse credibilidad a las explicaciones malintencionadas, que encuadran esas actividades en el contexto conspirativo postulando un largo plazo —muy largo, si fuera necesario— en el cual las actividades anticapitalistas del Estado son vistas como un prerrequisito funcional del capitalismo. "Es de notar que cualquier argumento puede ser acomodado a cualquier efecto, haciendo malabarismos con la escala del tiempo."[145] Finalmente, recordamos que una explicación funcional que no *especifique* la realimentación causal desde los efectos económicos de las

actividades del Estado hacia el mantenimiento de estas actividades es vacía y sin valor.

El Estado moderno y sus funcionarios tienen, en verdad, un deseo sincero de actuar como *intermediarios* neutrales entre los grupos conflictivos más bien que incentivar los intereses de un grupo particular. Sin embargo, la trampa está en el poder de decidir qué grupos participan en la negociación. Para el funcionario gubernamental moderno, forma parte de su segunda naturaleza el convencimiento de que sólo los grupos organizados pueden ser admitidos y que su organización, por así decirlo, es una condición para que se les tome en serio. Por lo tanto, concediendo que el Estado es aproximadamente neutral cuando media entre grupos organizados, la tendencia de la capacidad para organizarse se reflejará en el resultado de la mediación. Este es un caso de no toma de decisión, como el definido por Peter Bachrach y Morton Baratz: "Cuando los valores dominantes, las reglas aceptadas del juego, las relaciones de poder existentes en el interior de los grupos y los instrumentos de fuerza, por separado o de manera combinada, impiden realmente que algunos malestares se conviertan en cuestiones maduras que requieren decisiones, puede decirse que existe un estado de no toma de decisión."[146]

A las cuatro causas citadas de no toma de decisión —valores, normas, poder y fuerza— en este pasaje podemos añadir los atributos estructurales del grupo que constituyen el tema de las proposiciones (2) y (3). Ciertamente, esos atributos no están distribuidos de manera uniforme en todos los grupos que comparten las contradicciones en la sociedad y tienen una tendencia diferente, cuyo resultado es que los intereses del propietario y el administrador están mejor organizados que los intereses del trabajador, con un impacto mayor sobre las decisiones finales. No niego que los trabajadores en los países industrializados están altamente organizados, pero usualmente el horizonte de los sindicatos se detiene en la puerta de la fábrica. Los sindicatos están poco preocupados por los trabajadores como perpetradores y víctimas de la congestión y la polución, y no lo están por las condiciones de sus esposas e hijos. No hay, por ejemplo, organización que pueda impulsar las exigencias obvias de que el tiempo de viaje hasta el trabajo sea incluido en un número de horas de trabajo sujeto a una negociación colectiva. Los fabricantes de autos están mejor organizados que los trabajadores de la industria automotriz; los trabajadores, mejor que los dueños de los autos; los dueños de los autos, mejor que las víctimas de los autos, y las víctimas de los autos, mejor que las víctimas potenciales. A partir del hecho de que los intereses de la clase capitalista están sobrerrepresentados en la mayoría de las mediaciones se deduce que la neutralidad del Estado entre las diversas demandas postuladas o, incluso, una cierta tendencia a favor de los trabajadores (como en los países con gobiernos socialdemócratas) producirá un resultado final que es, también, asimétrico en la dirección de los intereses capitalistas.

El párrafo precedente indica "la naturaleza objetivamente capitalista del Estado moderno". Pienso que es posible avanzar un paso más y argumentar que esa representación asimétrica se mantiene por sus pro-

pios efectos. No sólo el conjunto de resultados finales está inclinado hacia el capitalismo sino que ello tiende a regenerar sus propias causas. Este circuito de realimentación surge del hecho de que un alto grado del factor de movilidad es parte importante de los intereses de la clase capitalista, de manera que en las decisiones finales tenderá a perpetuar y acelerar el alto ritmo del cambio de pertenencia, responsable en primer lugar de la decisión tendenciosa. La aparente inocencia del Estado resulta ser, de acuerdo con este análisis, parte de un círculo vicioso en el que la representación tendenciosa y el papel del *intermediario* imparcial generan una representación aún más tendenciosa.

(4) *La probabilidad de éxito en el intento de organizarse varía inversamente con la irreversibilidad de las contradicciones.*

Esto significa simplemente que cuando los actores finalmente se ponen de acuerdo sobre una estrategia cooperativa puede ya ser muy tarde. Además de los distintos espacios temporales ya discutidos, introducimos aquí el tiempo que se puede permitir que pase antes que la situación esté irreversiblemente deteriorada. Hay un tiempo en el que la población de una especie amenazada cae por debajo del mínimo requerido para mantener la reproducción y, si esto ocurre antes que la solidaridad se haya desarrollado, ninguna cantidad de cooperación posterior puede recrear a la especie extinta. En otros casos, la irreversibilidad es un poco menos extrema. A un costo muy alto, puede hacerse que los desiertos florezcan o restaurarse la vida de los lagos destruidos. Durante miles de años, esta contrafinalidad irreversible ha sido el principal factor del cambio social. La historia de China o de los países mediterráneos muestra que la erosión y la deforestación son los primeros agentes de cambio. Estos son casos de degradación irreversible, pero la acumulación irreversible es igualmente importante. Con algunas importantes excepciones, el *conocimiento*, una vez obtenido, no puede borrarse. Esto también es así para el conocimiento tecnológico, lo cual explica por qué Marx —en contraste con Keynes— no pensaba que el capitalismo podía superar las contradicciones del sistema. La caída de las tasa de ganancia, la principal contradicción que vio Marx en el modo de producción capitalista, estaba ligada al progreso técnico de manera tal, que era virtualmente irreversible. La teoría keynesiana de las crisis en la demanda no tiene implicancia de irreversibilidad y, en verdad, las crisis pueden ser revertidas por medio de la adecuada intervención gubernamental; el multiplicador funciona en ambos sentidos. Sin argumentación, agrego que, en mi opinión, sólo a través de esta noción de irreversibilidad, con los aspectos gemelos de degradación y acumulación irreversibles, puede dársele un significado preciso a la palabra "materialista", en la frase "la concepción materialista de la historia".

Por consiguiente, no exijo que las proposiciones (2) a (4) agoten el conjunto de variables que puede explicar el éxito o fracaso de los intentos de organización. Entre los factores discutidos por otros autores, el siguien-

te puede mencionarse brevemente. Granovetter enfatiza la estructura sociométrica del grupo en cuestión, destacando especialmente los débiles vínculos que sirven para acercar a los miembros a través de un tercero ("puentes").[147] North y Davis mencionan, entre otras cosas, la ganancia potencial por organizarse y el grado de incertidumbre asociado al resultado. También analizan un problema brevemente mencionado en el capítulo 3, el de que la probabilidad de acción colectiva exitosa es mayor "a mayor igualdad de distribución de las ganancias potenciales en [el] grupo".[148] Podrían haber agregado que este problema es especialmente agudo si el juego cooperativo en cuestión se resuelve a lo largo de líneas conocidas como la "solución de Nash"[149] en la que los beneficios de la organización son mayores, en primer lugar, para aquellos miembros que estaban mejor. En la discusión anterior, me he alejado deliberadamente de este problema, y he supuesto que todos los miembros del grupo generador de contradicciones pueden ganar la misma cantidad si se superan las contradicciones. Kenneth Arrow ha enfatizado la estructura de información del grupo, como una variable importante en la organización, observando que "El concepto de interés e identificación de clase puede estar relacionado con la facilidad de comunicación entre individuos con experiencias de vida similares".[150]

Cuando he destacado la distancia espacial, los ritmos de cambio de pertenencia y la irreversibilidad como variables fundamentales lo hice porque todas comprenden *el tiempo* de una manera que me parece esencial. (Es evidente que la distancia espacial tiene implicancias en el tiempo, a partir de que éste es necesario para desarrollar confianza y solidaridad en un grupo dado; mucho dependerá, entonces, de la distancia comunicativa entre sus miembros.) Es necesario que sincronice un número de procesos diversos para superar las contradicciones. El tiempo requerido para que surja la solidaridad de clase o grupo no puede exceder el tiempo de pertenencia al grupo ni el tiempo a partir del cual las contradicciones son irreversibles. Poner a las consideraciones temporales en el centro del análisis también permite hacer una nueva interpretación directamente *causal* de este frecuente dicho: *Plus ça change, plus c'est la même chose.* A mayor rapidez de cambio *dentro* de la estructura, menor probabilidad de cambio *de* la estructura. Es corriente en gran parte del pensamiento científico moderno la idea común de que las estructuras se mantienen a sí mismas por medio del flujo de materiales. La noción de *estado estable* —tan diferente a la noción de equilibrio— está extendida en la química, la biología y muchas de las ciencias sociales. Sin embargo, nos inclinamos a pensar que la estructura se mantiene a sí misma *a pesar de* los elementos cambiantes que la componen. En esta perspectiva, sugiero que es mejor la postura contraria: que un rápido metabolismo es una condición para el mantenimiento del sistema.

Las últimas observaciones están propuestas solamente para ser aplicadas a los sistemas *suboptimales.* El lector debe recordar que aquí estamos tratando con una teoría *dual* del cambio social. En la medida en que los fenómenos discutidos en las proposiciones (2) a (4) impidan la supera-

ción de la *contrafinalidad*, tendríamos más bien que invertir el argumento y decir que mucho cambio *dentro* de la estructura impide que los actores impidan un cambio *de* la estructura. Dependiendo del tipo de contradicción comprendido, las condiciones estructurales para la organización política pueden funcionar tanto como obstáculos y agentes para el cambio.

APÉNDICE 1 DEL CAPITULO 5

Algunas nociones y problemas de la teoría de juegos

Este apéndice tiene dos propósitos, uno elemental, el otro más complejo. En primer lugar, le ofrece al lector no familiarizado con la teoría de juegos los conceptos básicos necesarios para comprender los pasajes que anteceden sobre la suboptimalidad y la contrafinalidad. En segundo lugar, indica un vacío en el razonamiento de la última parte y esboza algunos métodos para llenarlo. Ese vacío se relaciona *grosso modo* con el mecanismo preciso de la transición de la conducta no cooperativa a la cooperativa. La teoría encerrada en las proposiciones (2)-(4) del último subcapítulo es en cierto sentido una "teoría de caja negra", porque sólo dice que dadas ciertas condiciones, el estado inicial de no cooperación tiende a transformarse en un estado final de cooperación, sin aclarar los detalles psicológicos del proceso. No pretendo ser capaz de resolver el problema, pero el lector tal vez quiera saber algo más sobre las soluciones y discusiones en la literatura especializada.

Un *juego*[151] incluye una cantidad de *actores*, cada uno de los cuales puede optar entre varias *estrategias*. Para cada combinación de estrategias —una por cada actor— las reglas del juego definen un *resultado* para cada actor. Así, en el caso general, la elección de *todos* define el resultado de *cada uno*. Los actores pueden definir sus estrategias de manera *cooperativa* o *no cooperativa*. La elección de la estrategia dependerá también de la *información* que se posee: sobre los resultados propios y ajenos, y sobre la información que poseen los demás jugadores. En ciertas condiciones, el juego también tendrá una *solución*, vale decir, un conjunto único y predecible de estrategias que elegirán los jugadores, o un conjunto único y predecible de desenlaces (es decir, desenlaces para los jugadores). Este es apenas el esqueleto de la teoría de juegos; veamos algo de la sustancia.

En aras de la sencillez, supondremos un juego con sólo dos actores. Puede parecer una hipótesis un tanto drástica en un estudio sobre la validez de la teoría de los juegos en las ciencias *sociales*, pero en verdad no es tan grave como parece. Porque, en primer lugar, muchos de los juegos de dos personas analizados en detalle más adelante se aplican fácilmente al caso de *n* personas.[152] Además, en segundo lugar, a muchos juegos entre dos personas se los puede interpretar como juegos entre uno mismo y " todos los demás". No obstante, sepa el lector que esta limitación excluye la parte de la teoría de los juegos potencialmente más pertinente para las ciencias

sociales, a saber, la *teoría de las coaliciones* en juegos cooperativos de n personas. Esta teoría ya cumple una función muy fecunda en la economía, y sin duda más adelante tendrá aplicaciones profundas en sociología. No obstante, el estado actual de los conocimientos justifica por ahora esta exclusión. También se aplica la hipótesis, menos drástica que la anterior, de que cada actor puede elegir entre un número finito de estrategias, generalmente dos. Por último, nos referiremos a juegos de suma variable, excluyendo los de suma cero. Lo que define a estos últimos es que la suma de los resultados de los jugadores es constante (y por redefinición del juego se puede dar a esta constante el valor cero), mientras que en los primeros la suma de los resultados obtenidos por todos los jugadores depende de las estrategias elegidas. Los juegos de suma cero se caracterizan por el conflicto puro, mientras que los que interesan aquí combinan el conflicto y la cooperación.

El concepto de *elección de estrategia* puede adquirir gran complejidad. El caso más sencillo es aquel en que cada jugador individual elige una estrategia propia y predecible entre sus estrategias puras, es decir, pertenecientes al conjunto dado (que para el caso se considera finito). Un caso mucho más complejo es aquel en que cada actor elige individualmente, pero al azar, de acuerdo con una cierta distribución de probabilidades sobre el conjunto de estrategias puras. Esta distribución se llama *estrategia mixta*. Si un juego tiene una solución en estrategias mixtas, significa que las distribuciones de probabilidad que rigen la elección de estrategias puras son individualmente predecibles, aun cuando las estrategias puras en sí no lo sean. Antes de dar un ejemplo de un juego que no tiene solución con estrategias puras, pero sí cuando se admiten estrategias mixtas, es necesario aclarar la nomenclatura. Llámense los jugadores Hilera y Columna, que pueden optar por las estrategias h_i y c_i respectivamente. En forma de matriz, cada casillero representa una combinación de estrategias: el primer número del casillero indica el resultado obtenido por Hilera, el segundo por Columna. Con estas convenciones, se advierte fácilmente que la solución del juego siguiente (de suma cero) requiere estrategias mixtas:

	c_1	c_2
h_1	3,-3	1,-1
h_2	2,-2	4,-4

Además de la extensión de estrategias puras a mixtas, ahora se introduce la extensión de la elección individual a la elección conjunta de estrategias. Es muy importante el concepto de *estrategias mixtas conjuntas*. Por ejemplo, si Hilera y Columna pueden elegir entre (h_1,h_2) y (c_1,c_2) respectivamente, tal vez opten por la estrategia mixta conjunta de (h_1,c_1) con una probabilidad de 1/2 y (h_2, c_2) con una probabilidad de 1/2. Este podría

ser un procedimiento útil si tanto (h_1,c_2) como (h_2,c_1) tienen consecuencias desastrosas, como se muestra más adelante en uno de los ejemplos. Si los dos actores optan por estrategias mixtas conjuntas, no sería posible evitar estos dos desenlaces y a la vez conservar los dos primeros. El uso de estrategias mixtas conjuntas es una característica crucial de los juegos cooperativos.

Otro concepto importante y complejo de la teoría de los juegos es el de *solución*. Limitándose al caso no cooperativo, el concepto se define de la siguiente manera. Primero se define el concepto de *punto de equilibrio*: es un par de estrategias puras o mixtas (h_i,c_j) tales que h_i es la mejor respuesta de Hilera a c_j y c_j es la mejor respuesta de Columna a h_i. Si el juego tiene un punto de equilibrio exacto, ésta es la solución. Si no hay punto de equilibrio, el juego no tiene solución. (Confróntese lo dicho en el último subcapítulo sobre la hiperinflación como ejemplo posible de un juego sin punto de equilibrio.) Si existen varios puntos de equilibrio, el juego tiene solución si y sólo si todos los puntos de equilibrio no dominados por otro punto de equilibrio son equivalentes e intercambiables. Esta afirmación presupone los siguientes conceptos. Puntos de equilibrio *equivalentes* son aquellos que dan los mismos resultados a todos los jugadores. Los equilibrios (h_i,c_j) y (h_m,c_n) son *intercambiables* si (h_i,c_n) y (h_m,c_j) también son puntos de equilibrio. Que (h_i,c_j) *domina* (h_m,c_n) significa que el primero es Pareto-superior al segundo, de manera que en el primer caso los dos actores consiguen al menos lo mismo y uno de ellos un poco más. Más adelante, en algunos ejemplos numéricos, se exploran las implicancias de esta definición del concepto de solución. Baste decir por el momento que trata de aprehender el concepto de lo que harían actores racionales en un juego no cooperativo si contaran con información perfecta.[153]

Los cinco juegos siguientes sirven para destacar algunos de los problemas que surgen en juegos de suma variable:

Dilema del Prisionero

	c_1	c_2
h_1	5,5	- 10, 10
h_2	10,-10	0, 0

Juego de Certeza

	c_1	c_2
h_1	10, 10	-10,5
h_2	5,-10	0, 0

Batalla de los Sexos

	c_1	c_2
h_1	2, 1	-1, -1
h_2	-1,-1	1, 2

Gallina (I)

	c_1	c_2
h_1	1, 1	- 2,2
h_2	2, -2	-5, -5

Gallina (II)

	c_1	c_2
h_1	1, 1	-2, 6
h_2	6, - 2	-5, -5

Estos juegos han sido objeto de una gran cantidad de estudios.[154] El más conocido es el Dilema del Prisionero que, como se ha dicho, está en la base de importantes situaciones sociales. La solución es (h_2,c_2) —es el único punto de equilibrio— y es evidentemente subóptima. Obsérvese que en estos casos las segundas estrategias de Hilera y Columna son las *dominantes*. Para Hilera, h_2 es la mejor, independientemente de lo que haga Columna, y asimismo, c_2 es la mejor tanto contra h_1 como h_2. Los demás juegos tienen varios puntos de equilibrio y carecen de estrategias dominantes. El Juego de Certeza tiene los equilibrios (h_1,c_1) y (h_2,c_2), de los cuales el primero debe ser la solución, puesto que domina al segundo. (Adviértase la diferencia entre una estrategia dominante y un conjunto de estrategias que domina a otro conjunto.) El Juego de Certeza también puede generar una situación subóptima, por ejemplo, si los jugadores carecen de la información necesaria para converger sobre la solución. Si en este juego cada jugador conoce sólo sus propios resultados pero sospecha que la estructura de resultados de su contrincante es la del Dilema del Prisionero, o que el otro sospecha que *su* estructura lo es, cada uno esperará del otro que elija su segunda estrategia y por consiguiente él mismo deberá hacerlo. Así, la suboptimalidad puede resultar de la falta de información (como en el Juego de

Certeza) o estar arraigada profundamente en el modo no cooperativo de la elección en sí misma (como en el Dilema del Prisionero).

La Batalla de los Sexos y el de la Gallina son juegos sin solución, aunque se permitan las estrategias mixtas. Tienen varios equilibrios: (h_1,c_1) y (h_1,c_2) en el primero; (h_1,c_2) y (h_2,c_1) en el segundo. Estos equilibrios no son dominantes, equivalentes ni intercambiables.[155] Lo que *sucederá* en la Batalla de los Sexos o en la Gallina no es asunto de la racionalidad sino de la psicología, salvo que estemos dispuestos a aceptar una extensión adicional del marco. La introducción del concepto de estrategias mixtas obedeció a la existencia de juegos sin solución en estrategias puras; asimismo, la existencia de juegos no cooperativos sin solución en estrategias mixtas puede motivar la introducción de soluciones cooperativas. Existen varias soluciones conceptuales para juegos cooperativos de dos personas, pero la falta de espacio no permite entrar en detalles. Aquí nos concentraremos en la Batalla de los Sexos y las dos versiones de la Gallina para llegar a lo que considero son las ideas esenciales. En Gallina (I) se advierte intuitivamente que los dos jugadores que cooperan se pondrán de acuerdo sobre el par estratégico (h_1,c_1). En cambio, en Gallina (II), la ventaja que otorga el uso unilateral de la segunda estrategia es tan grande que se buscará una solución cooperativa diferente, a saber, la estrategia mixta conjunta donde los jugadores tiran la moneda para optar por (h_1,c_2) o (h_2,c_1), lo que da a cada jugador un resultado esperado 2. (De paso, la diferencia entre Gallina (I) y Gallina (II) muestra que, si se miden los resultados sólo en unidades ordinales, como a veces se nos propone, la solución cooperativa "natural" podría ser indeterminada.)[156] En la Batalla de los Sexos la solución cooperativa también es una estrategia mixta conjunta, en que las probabilidades están repartidas por igual entre los dos equilibrios.

Con las observaciones precedentes concluye la parte elemental de este apéndice y pasamos a conceptos más complejos (o en todo caso, menos corrientes). Como ya se ha dicho, el problema es cómo analizar el paso de la conducta no cooperativa a la cooperativa. Por empezar, adviértase que este problema admite dos interpretaciones distintas. Primero, cabe preguntarse en qué condiciones se observará el paso de la *estrategia* no cooperativa a la cooperativa, por ejemplo en el Dilema del Prisionero. Segundo, cabe preguntarse en qué condiciones se verá el paso del *modo de decisión* no cooperativo al cooperativo, por ejemplo en la Batalla de los Sexos. Como se verá más adelante, el primero de estos problemas incluye nuevas ambigüedades, pero se demostrará que los subproblemas derivados admiten respuestas bastante concretas. En cambio, el segundo problema parece más difícil y no nos extenderemos sobre él.

En el Dilema del Prisionero es habitual calificar las estrategias h_1 y c_1 de *cooperativas* o *altruistas*, y las h_2 y c_2 de *no cooperativas* o *egoístas*. Se ha analizado el paso del altruismo al egoísmo de acuerdo con tres lineamientos distintos. En primer lugar, se puede adoptar la imagen causal del hombre en lugar de la intencional (véase el apéndice 2) y analizar las condiciones

en las que sería dable esperar desviaciones de la conducta racional. A este título corresponde buena parte del trabajo realizado por los psicológicos sociales con el Dilema del Prisionero. Cabe mencionar sobre todo la obra empírica y teórica asociada con el nombre de Anatol Rapoport.[157] Una muestra típica de este enfoque es el modelo de Asch-Cohen de las series de dilemas del prisionero.[158] Aquí se presupone que cada jugador se encontrará en alguno de los estados siguientes:

s_1: si se halla en este estado, usará la estrategia cooperativa en todos los juegos siguientes;

s_2: si se halla en este estado, usará la estrategia cooperativa en el juego siguiente, pero podría usar la estrategia no cooperativa en los subsiguientes;

s_3: en este estado usará la estrategia no cooperativa en el juego siguiente, pero podría pasar a la cooperativa en los subsiguientes;

s_4: en este estado, usará la estrategia no cooperativa en todos los juegos siguientes.

Además, suponemos que el proceso de transición es una cadena de Markov. A partir de estos presupuestos podemos deducir varias conclusiones relativas a la conducta esperada de los jugadores y verificarlas con las pruebas empíricas. Evidentemente, el rasgo característico de este enfoque es que no se visualiza a los jugadores como actores racionales sino como vehículos de procesos causales; como portadores de "estados" en lugar de "intenciones". Es difícil decir si los numerosos resultados interesantes obtenidos por medio de este enfoque y otros similares se aplican en casos de la vida real con muchas personas, y en todo caso no se hablará más sobre esto.

En segundo lugar, podemos afirmar que la práctica repetida de los juegos conduce a un cambio en la estructura de preferencias, de manera que el comportamiento altruista se torna racional. Esto significa simplemente que luego de jugar repetidamente, el juego ya no *sería* el Dilema del Prisionero, sino otro, como el Juego de Certeza. Si nos contentamos con observar el comportamiento de los actores, parecería como si eligieran la estrategia cooperativa del Dilema del Prisionero, pero pueden estar pasando cosas en sus cabezas que produzcan una situación diferente. Más concretamente, podemos suponer que, en la medida en que los actores comienzan a conocerse y sentir mutuo afecto, (i) obtienen alguna utilidad positiva del bienestar de los demás, (ii) obtienen alguna utilidad negativa de su utilización de una estrategia egoísta ("mala conciencia") y (iii) llegan a obtener alguna utilidad positiva de la igualdad de ingresos como tal. Si la estructura del resultado material (medido, digamos, en dinero) es un Dilema del Prisionero, la estructura real de preferencias puede ser muy diferente si refleja uno o más de estos mecanismos. La estructura real de preferencias puede ser tal, que haga de la cooperación la estrategia

dominante, o más probable, quizá, que la cooperación sea una solución sin ser una estrategia dominante (como lo es en el Juego de Certeza). Presumiblemente, las mismas condiciones —exploradas en las proposiciones (2) y (3) en la última sección— transforman el juego del Dilema del Prisionero y hacen que los actores puedan realizar una solución que no consista en estrategias dominantes. El afecto por los demás y la información acerca de ellos crece *pari passu*, o algo similar.

La aproximación explicada en este último párrafo tiene de todos modos la virtud de la simplicidad, aun si se dice que resuelve el dilema de una manera trivial. Una aproximación más ambiciosa sería argumentar que puede ser racional utilizar la estrategia cooperativa incluso si el juego es un Dilema del Prisionero, en sentido estricto. La primera línea de argumentación esbozada resolvió el dilema renunciando a la racionalidad, la segunda, cambiando el juego; llegamos ahora a la tercera aproximación, la cual rechaza cualquiera de estas salidas fáciles. Este enfoque existe en dos variantes principales, la teoría del "metajuego", propuesta por Nigel Howard,[159] y la teoría del "superjuego", propuesta por Martin Shubik y Michael Taylor.[160] Como pienso que no muchos estudiosos en este campo podrían aceptar que la teoría del metajuego tenga realmente poder resolutivo, sólo haré comentarios breves acerca del método del superjuego, que es mucho más prometedor. Una vez más, en el corazón del tema hay una secuencia del Dilema del Prisionero. Es bien sabido que, en una secuencia de juegos como éstos, el comportamiento no cooperativo será elegido en cada juego si (i) todos los juegos cuentan por igual (sin descontar tiempo del futuro), y (ii) si se fija el número de juegos por adelantado y esto lo conocen los jugadores.[161] Contrariamente, el comportamiento cooperativo puede ser la elección racional si (i) hay una preferencia por el presente, (ii) el número de juegos se fija con anterioridad, pero los jugadores no lo conocen, o (iii) el número de juegos es determinado por un proceso estocástico.[162] La primera de estas posibilidades se encuentra en el corazón del método de Shubik-Taylor. Michael Taylor, en particular, ha demostrado cómo ciertas condiciones harán que surjan *estrategias condicionalmente cooperativas*, como en la solución a un Dilema del Prisionero secuencial.

Para mayores detalles sobre el método de Taylor, tengo que remitir al lector al original. Aquí, sólo enfatizaré el hecho principal de que estas estrategias condicionalmente cooperativas no son estrategias dominantes, lo cual significa que es bastante dudosa la probabilidad de que se desarrollen en juegos de la vida real. El mismo Taylor es muy consciente de este punto:

> el requisito de un alto nivel de conciencia por parte de los cooperadores condicionales es "más probable" en un grupo pequeño de jugadores que en uno grande, y aún más en el tipo de comunidad pequeña donde las personas están en contacto y pueden observar el comportamiento de muchos de sus compañeros, y el cual es bastante estático, en el sentido de que no hay mucha movilidad hacia dentro y hacia fuera.[163]

Aunque concuerda con el razonamiento que respalda las proposiciones (2) y (3) de la última sección, esto quizás esté sujeto a la objeción de que el tipo de grupo que favorece la solución —una pequeña comunidad con poca movilidad— sea también el grupo que esté a favor de abolir el juego. Las personas pueden cambiar sus preferencias si interactúan de cerca, de manera de tornar irrelevante el hecho de que la interacción también permita una solución al dilema creado por viejas preferencias.

Pienso que la elección entre las tres aproximaciones al "problema de transición" es un problema de investigación, que puede hacerse de manera bastante directa. Se requerirá un duro trabajo para determinar cuál aproximación se adecua mejor a qué casos, pero no parece haber ningún obstáculo conceptual inherente al análisis. Por otra parte, dichos obstáculos se presentan tan pronto como formulamos el problema de una transición de la decisión no cooperativa a la cooperativa. J.F.Nash[164] ha sugerido que la decisión de cooperar debería verse como movimiento de un juego más amplio de no cooperación, porque este último es algo más fundamental. Me parece que este punto de vista es sólido y surge del principio del individualismo metodológico, pero también estoy de acuerdo con los críticos que señalan que carece de poder persuasivo mientras no sea explicado en detalle.[165] La confusión acerca de dos aspectos de la solución del propio Nash al problema de la negociación es un problema típico. La parte más conocida de esta solución nos dice cómo encontrar la solución cooperativa, *dados los resultados del no intercambio* o *status quo*. Si conocemos los resultados que obtendrían los jugadores que no se comprometieran en ninguna negociación, podríamos argumentar que la solución de cooperación está determinada de manera única. La solución, sin embargo, tiene otro aspecto, que es, la determinación del punto de no intercambio. Por medio de amenazas, cada jugador luchará por obtener un punto de no intercambio que haga que la solución cooperativa sea lo más favorable posible para él. Parece muy difícil conceptualmente entender si el juego no cooperativo (utilizando las amenazas como estrategias) o el cooperativo (utilizando conjuntamente estrategias mixtas) es el más básico. Hasta que no haya una mayor clarificación de estos temas y los que se le relacionen, se debería ser cauto en la utilización de la teoría de juegos cooperativo con fines empíricos.

APÉNDICE 2 DEL CAPITULO 5

Causalidad e intencionalidad: tres modelos de hombre

En este apéndice intento juntar algunos de los estándares de los capítulos 4 y 5, explorando la relación entre las contradicciones, la intencionalidad, la causalidad, el pensamiento estratégico y la libertad. El postulado básico del cual parto es: *la meta de las ciencias sociales es la liberación del*

hombre. Tendrían que permitirle liberarse de las fuerzas causales que forman y pervierten sus intenciones; posibilitarle realizar sin obstáculos las metas libremente elegidas. En verdad, éste es un objetivo difícil, y una de las conclusiones del análisis será que no puede alcanzarse. Sin embargo, pienso que ésta debería ser la idea que guíe siempre al científico social.

Desde mi punto de vista, hay cuatro obstáculos para la realización de esta idea. En primer lugar, como lo establecimos en el capítulo 4, el hombre se fija metas inconsistentes o contradictorias. En segundo lugar, el individuo está atravesado por procesos causales que escapan a su dominio y no entiende. Pertenece a este problema el cambio endógeno de preferencias, discutido en los capítulos 2 y 4. En tercer lugar, un grupo de personas, cada una de las cuales tiene preferencias consistentes y objetivos libremente escogidos, puede, sin embargo, no tener la posibilidad de obtener lo buscado, por el problema de la contrafinalidad. En cuarto lugar, hemos visto que la transición al modo estratégico de pensamiento no asegura la realización de las metas, porque el juego que se sigue puede ser un juego sin solución. Es tentador agregar las soluciones suboptimales como un quinto obstáculo, pero esto traería consigo un problema *sustantivo*, que va más allá de los problemas estrictamente *formales* inherentes a los primeros cuatro obstáculos. Tanto en el nivel colectivo como individual, la libertad puede ser utilizada para finalidades malas o buenas; así como el suicidio puede presumiblemente ser un acto que cumple con los requisitos formales de una acción libre, también puede serlo la elección de una estrategia no cooperativa en el Dilema del Prisionero.

Por lo tanto, sostengo que los obstáculos a la libertad pueden resumirse como contradicciones mentales, causalidad subintencional, contradicciones sociales (= causalidad supraintencional) y juegos sin soluciones. Para nuestros objetivos, esto puede simplificarse de manera de que queden sólo dos clases de obstáculos. Las contradicciones de la mente pueden subsumirse bajo el título de causalidad subintencional, porque en dichos casos tiene que haber algún mecanismo causal que determine de cuál de los deseos contradictorios prevalece en una situación dada. También pienso que los juegos sin una solución están cubiertos por la noción de causalidad supraintencional, porque el resultado será determinado por la (impredecible) interacción causal entre las acciones generadas por las suposiciones arracionales acerca del comportamiento de los demás.

Por lo tanto, nos queda el clásico problema de la causalidad versus la intencionalidad en el análisis de las acciones humanas, o de explicación versus comprensión. No me detendré en las espinosas cuestiones filosóficas, excepto para hacer las siguientes observaciones. En primer lugar, argumento a favor de una tricotomía más bien que de una dicotomía: en vez de la simple diferenciación entre intencionalidad y causalidad, propongo dividir a ésta en causalidad intencional y causalidad supraintencional. (Es útil hacer una diferenciación similar en biología, donde las excepciones a la afirmación "todo tiene una función" pueden encontrarse tanto en el nivel subfuncional como suprafuncional.)[166] En segundo lugar, me parece que las explicaciones intencionales de las acciones humanas no pueden reducirse

a explicaciones causales, en el sentido de una reducción al lenguaje neurofisiológico de las causas inmediatas,[167] ni tampoco en el de una reducción al lenguaje evolutivo de las causas últimas.[168] En tercer lugar, creo que la afirmación precedente es totalmente compatible con dos afirmaciones que *prima facie* parecen contradecirla. Que algunas formas de comportamiento humano sólo pueden ser analizadas causalmente y no comprendidas intencionalmente, y que todas las formas de comportamiento humano, en principio, pueden explicarse sobre una base puramente causal. Como se menciona en el capítulo 1, la conclusión del materialismo y el determinismo a partir del reduccionismo es una falacia similar a la conclusión, a partir de "todo tiene una causa", de "hay algo que es la causa de todo".

Implícito en esto está el hecho de que en varios contextos uno de tres modelos diferentes de hombre será el apropiado. Los llamaré imágenes del hombre causal, paramétrica intencional y estratégica intencional. El hombre estratégico intencional es el actor de la teoría de juegos que sabe que su medio ambiente está conformado por otros actores estratégicos, y que él es parte del medio ambiente de ellos. Por lo tanto, podrá realizar sus metas sin ningún peligro de contrafinalidad. *La acción estratégica es la encarnación de la libertad colectiva*, la racionalidad totalmente transparente, que se incorpora a las expectativas de los demás actores racionales para converger en un curso de acción predecible. En contraste, el comportamiento paramétrico intencional surge de un estado mental más confuso. El actor paramétrico intencional supone que es libre de adaptarse óptimamente a un ambiente constante o paramétrico. Como hemos visto, esta presunción es, en sí misma, bastante consistente, pero, si todos los actores la sostienen simultáneamente, genera contrafinalidad y en esa medida constituye un obstáculo para la libertad. Me parece muy claro que las ciencias sociales pueden ir muy lejos en la tarea de lograr una transición de una sociedad compuesta por actores paramétricos intencionales a otra de estratégicos intencionales. Repito que esta libertad colectiva no garantiza la racionalidad colectiva, en el sentido de que se den situaciones Pareto-optimales; el Dilema del Prisionero exhibe la posibilidad de la autodestrucción libremente elegida. Sin embargo, no creo que la transición al pensamiento estratégico pueda hacer que la situación sea peor que antes y, en muchos casos —tal como en el Juego de Certeza— la transición puede ser suficiente para que aparezca un comportamiento colectivo racional.

El problema de la causalidad subintencional es más oscuro. Para nuestros propósitos, puede dividirse en varios subproblemas, todos con implicaciones negativas para la libertad. Ellos son los problemas de las metas contradictorias, el del comportamiento no dirigido hacia una meta y el de la formación de metas. Se ha dicho ya lo suficiente sobre el primero, por lo que aquí seré breve. En este caso, la dificultad es que tenemos *demasiadas intenciones*. Incluso si las ciencias sociales, la psicología y diversas formas de terapia en particular pueden ser, presumiblemente, de ayuda para resolver los nudos mentales señalados en el capítulo 4, la

existencia de varias intenciones inconsistentes vuelve difícil decir con exactitud a *quién* se está ayudando. ¿Debería liberarse el amo paternalista del sádico o viceversa? Y ¿puede una persona como esa liberarse a "sí misma"? Aquí parece haber una dificultad conceptual que por el momento se me escapa.

En el segundo problema, la dificultad más bien es que no tenemos ninguna intención. Creo que en este punto son necesarios algunos ejemplos para mostrar que la acción puede no estar dirigida a una meta. En verdad, hay dos formas diferentes en que puede negarse la realidad de la elección intencional entre alternativas. Para explicar la diferencia, tenemos, primero, que observar que cualquier acción realizada puede ser vista como el resultado de dos procesos de filtro sucesivos. En primer lugar, hay restricciones objetivas que disminuyen el conjunto de cursos de acción abstractamente posibles a un pequeño subconjunto, *el conjunto factible*. En segundo lugar, tenemos que postular algún mecanismo para explicar por qué se realiza una acción en vez de otra del conjunto factible. En la teoría de la elección racional, ese mecanismo es la maximización de alguna función objetiva. En el modelo de dos pasos surge de inmediato que la realidad de la elección racional puede ser negada, *o bien* postulando que el conjunto de restricciones, efectivamente, define un conjunto factible de sólo un miembro, o, de manera más realista, un conjunto tan pequeño que hace imposible distinguir sus miembros, *o bien* argumentando que la realización de una acción en vez de otra del conjunto factible no se hace por medio de la elección racional. Pienso que algunos de los escritos de la escuela estructuralista francesa pueden ser interpretados de acuerdo con la primera posibilidad, como también algunas de las argumentaciones de Marx que conciernen a los coeficientes fijos de producción y consumo.[169] Sin embargo, no me detendré en esto porque mi interés principal para nuestros fines es la segunda posibilidad. Gary Becker[170] ha diferenciado aquí los dos subcasos. El comportamiento irracional puede significar o bien que la realización de un miembro del conjunto factible esté gobernada por la *tradición* o bien que ocurra, más o menos, por *azar*. Veamos estos dos casos.

Incluso si, como lo implican las anteriores observaciones, hay diversas alternativas al modelo de elección racional, la alternativa más frecuentemente citada es probablemente el modelo orientado por normas. De acuerdo con el método de elección racional, la acción realizada cambiará típicamente si hay un cambio en las restricciones que definen al conjunto factible, mientras que el modelo orientado por normas niega que éste sea el caso. Como observó Becker, aquí hay una inconsistencia seria porque, si las restricciones cambian de manera que el viejo curso de acción queda por fuera del nuevo conjunto factible, *tiene* que haber algún cambio en el comportamiento observado. La ocupación no puede ser heredada de padre a hijo si el progreso técnico transforma en obsoleta la ocupación del padre; los hábitos de consumo tienen que cambiar si el anterior punto de

consumo queda por encima del nuevo presupuesto. Además de este problema conceptual, se pueden hacer muchas objeciones empíricas pertinentes al modelo puro orientado por la norma. Un buen ejemplo es la discusión de la "economía del delito". William Cobb y Gordon Tullock,[171] entre otros, han establecido la diferencia entre dos hipótesis que intentan explicar por qué algunas personas se hacen criminales: la hipótesis "económica", que supone la maximización de la utilidad esperada, y la hipótesis "sociológica" (también llamada hipótesis de la "enfermedad"), que liga las actividades delictivas a normas específicas de la subcultura, que más o menos inexorablemente empujan al delito a algunos individuos. La primera explicación es intencional y relaciona una acción con sus consecuencias esperadas en el futuro; la segunda es causal, articulando la acción a las condiciones del pasado.

Una consecuencia empírica de la hipótesis sociológica es que la elección de una carrera criminal no debería verse significativamente afectada por los cambios en la probabilidad del fallo condenatorio o por la severidad del castigo. Varios estudios han encontrado que esto es falso. De acuerdo con Tullock,[172] muchos de los sociólogos que comenzaron a probar la hipótesis de que el castigo no detiene el crimen terminaron encontrando que tiene un efecto de contención. Además de esta objeción empírica y del argumento conceptual citado en el último párrafo, está el siguiente punto teórico, que también habla en contra de la hipótesis sociológica. En verdad, tendría que aceptarse el postulado de que las normas son específicas de la subcultura, pero esto no quiere decir que ellas determinen *directamente* el curso de acción. La articulación es más bien indirecta: las normas determinan las preferencias; las preferencias + el conjunto factible determinan las acciones. Aquí vemos el problema de la formación de la preferencia, al cual volveré. Sin embargo, quisiera señalar que entre el modelo puro de elección racional y el modelo determinado por normas puede haber algunos esquemas intermedios. En tanto que la acomodación a los cambios en el sistema de recompensas no es instantáneo, el retraso puede deberse a una inercia psíquica, que es un fenómeno puramente causal. En algunas situaciones, sé que la inercia puede ser una estrategia racional y que, en otras, quizá, puede explicarse por los "costos psíquicos" del acomodamiento. Sin embargo, negaría que *todos* los casos de comportamiento aparentemente irracional puedan comprenderse como residuos de alguna racionalidad de orden superior.[173]

Yendo al segundo subcaso de Becker, a primera vista parecería que todos los modelos probabilísticos de acciones individuales tienen que apoyarse en la imagen causal del hombre más bien que en la intencional. Por ejemplo, los modelos de movilidad social parecen suponer que el hombre es sólo un vehículo de probabilidades de transición; más que encontrarse *atraído* por alguna ocupación es *empujado* hacia ella, con una probabilidad que depende de su historia ocupacional. Recientemente, sin embargo, ha habido varios intentos de llevar la intencionalidad y la racionalidad a los modelos de movilidad.[174] El más prometedor; parece ser el de

Raymond Boudon. El lector recordará de la anterior discusión sobre el trabajo de Boudon que su manera de reconciliar la intencionalidad y la probabilidad utilizaba la noción de estrategia mixta. Pero podría haber otra forma de reconciliar a ambas, que surge del enfoque de dos pasos de la acción, antes esbozado: ubicando el elemento probabilístico en el nivel del conjunto factible más bien que en el de la elección misma. O sea, si vamos a predecir la elección de un actor racional que tendrá que escoger dentro de un conjunto factible, sujeto a variación de acuerdo con alguna de distribución de probabilidad, entonces nuestra predicción también tiene que ser probabilística. (En principio, la probabilidad podría entrar en ambos niveles simultáneamente.) No quiero argumentar que por medio de uno de esos mecanismos *todas* las explicaciones probabilísticas pueden ser transformadas en explicaciones intencionales. Hay ciertos procesos causales trabajando en la mente, que toman un carácter estocástico, en el nivel del comportamiento observado; el modelo de Asch-Cohen-Rapoport para el Dilema del Prisionero, citado en el apéndice 1, es sólo uno de los muchos ejemplos de dichos procesos en la psicología experimental. Sin embargo, quiero reiterar que los científicos sociales tendrían que estudiar tales procesos, con el propósito de la liberación más que de la manipulación.

Concluyo este apéndice con algunas observaciones acerca del tercero de los subproblemas inherentes a la causalidad subintencional, la cuestión de la formación de preferencias y el cambio endógeno de preferencias. En este punto, la noción de *libertad* requiere una discusión menos superficial que la anterior. Me parece que la noción que está detrás de la idea kantiana de libertad es que el hombre, para ser libre, de alguna manera tendría que estar en posibilidad de *elegir él mismo*, no sólo en el débil sentido de actuar de acuerdo con preferencias consistentes de cualquier nivel, sino también en el sentido más fuerte de haber elegido esas preferencias. Esto es análogo al problema propuesto por Sidney Winter en sus profundas discusiones acerca de la maximización de la ganancia. Su contribución puede verse como el cuarto y decisivo paso en una serie de intentos por comprender qué significa realmente la racionalidad en los negocios. El primer intento era el modelo estándar de maximización de ganancias, presumiendo información sin costo y ningún costo de transacción. El segundo fue el modelo de "satisfacción" de Herbert Simon, quien argumentó que en la práctica real los hombres de negocios buscan un nivel satisfactorio, más que máximo, de ganancias.[175] El tercer paso fue una réplica de la Escuela de Maximización, diciendo que la satisfacción emerge como una variedad de la maximización, una vez que se toman en cuenta los costos de adquirir y evaluar información.[176] En un trabajo sorprendentemente ignorado, Winter argumentó que esa réplica creaba una regresión infinita porque ¿cómo se resuelve el problema de crear una cantidad óptima de información? La "elección de una estructura que maximice la información requiere ella misma información y no es claro cómo el aspirante a maximizar las ganancias adquiere esa información o qué garantiza que no pague un precio excesivo por ella".[177] En un trabajo posterior, Winter vuelve en estos términos sobre el problema: una maximización verdadera sería

una "optimización cuyo alcance cubre todas las consideraciones, incluyendo sus propios costos".[178]

El problema de la libertad es análogo al de la maximización de ganancias. En primer lugar, surge la misma paradoja de información. Aristóteles dijo que se puede ser responsable de la propia ignorancia,[179] pero ¿cómo puede alguien saber cuánto tendría uno que conocer? En segundo lugar, y éste es el problema más profundo, el hecho de que se elija de acuerdo con las preferencias no significa que se elija lo que se prefiere. Si existieran algunas *preferencias de punto fijo* que pudieran, por así decirlo, justificarse a sí mismas, ello podría resolver el problema. Pero, primero, hasta que la posibilidad de dichas preferencias sea exhibida en detalle no sé si la noción es realmente significativa y, segundo, podrían existir varios puntos fijos y la elección entre ellos sería difícil de explicar racionalmente. (Este problema corresponde en parte a algunas bien conocidas dificultades cartesianas de la elección de un mundo, entre muchos mundos posibles por parte de Dios.)[180] Queda pendiente una solución a estos problemas y por lo tanto pienso que la cuestión de la formación de preferencias y de los cambios endógenos de preferencias es el mayor obstáculo a la libertad total. Puede haber un proceso interminable de liberación, porque en cualquier nivel dado de, digamos, cambio endógeno de preferencias, es posible moverse un nivel hacia arriba y aprovechar la causalidad psíquica, con propósitos estratégicos propios en este nivel superior. Cada proceso como éste tiene finalmente que detenerse en algún nivel y, en este sentido, la libertad es imposible. Pero no hay un nivel en el que este proceso tenga que detenerse. Y en este sentido, la libertad, como idea regulativa, es una noción verdaderamente significativa.

Notas

1. Robinson (1956, pág.78); véase también Keynes (1971, págs.143 a 145). Como parte de mi intención es ligar el análisis de las contradicciones con el trabajo de Hegel y Marx, daré una larga cita de Marx para demostrar que él era muy consciente de este problema, incluso si no lo vio como la paradoja esencial del capitalismo. "Ningún economista negará que si los trabajadores *generalmente*, esto es, como *trabajadores* (lo que el trabajador como individuo hace o puede hacer por fuera de su género sólo puede existir como *excepción*, no como *regla*, porque no es inherente al carácter de la relación misma), actúan de acuerdo con [la demanda de ahorrar] como una *regla* (aparte del daño que harían al consumo general —la pérdida sería enorme— y, por lo tanto, también, para la producción, por lo tanto, también, para el volumen del intercambio que podrían hacer con el capital y, por lo tanto, para ellos mismos como trabajadores), entonces, los trabajadores estarían empleando medios que contradicen absolutamente sus propósitos... Cada capitalista exige que sus propios trabajadores ahorren, pero sólo los *suyos propios*, porque para él, ellos son trabajadores; pero, de ninguna manera, el restante *mundo de los trabajadores*, porque para él, ellos son consumidores" (Marx, 1857-1858, págs.285,287). La razón por la que Marx imputa al capitalista individual un deseo de que sus trabajadores ahorren es que esto le permitiría bajar los salarios durante una recesión.

2. Lazarsfeld y Menzel (1969, pág.507).

3. *Ibíd.*

4. Las excepciones son Vendler (1967) y Geach (1972, págs. 1-12). Las investigaciones de Hamblin (1970) y Mackie (1976) no se refieren a la falacia.

5. Ullman-Margalit (1976); véase también Sobel (1967) para algunos de los mismos puntos, aunque menos claramente establecidos.

6. Para este argumento véase Singer (1963), Lyons (1965) y Sobel (1967).

7. P.y T.Ehrenfest (1959,pág. 4).

8. Geach (1972, pág.1).

9. Véase Guilbaud (1952) para un brillante análisis de una controversia de Quetelet-Cournot acerca de esta noción. Brevemente, la idea es que pueda ser físicamente imposible para cualquier individuo aislado tener el valor promedio de *todas* las variables simultáneamente. Por ejemplo, si hay creciente retorno en escala, entonces, en una población de empresas eficientes no habrá ninguna empresa que tenga a la vez el tamaño promedio y el ingreso promedio.

10. Winch (1958, pág.16).

11. Hume (1748, sección IV, parte II).

12. Wittgenstein (1953, pág. 345). Refiriéndose a la discusión sobre los diversos tipos la universalizabilidad, vemos que Wittgenstein está desarrollando dos de ellas en una sola inferencia $(\exists x)\ (M(Fx)) \supset M\ (\forall x)(Fx)$.

13. Sin embargo, la contradicción entra de manera diferente. Como se mencionó antes, Kant no nos dice que evitemos una máxima si no podemos *imaginar* que pueda ser transformada en una ley general, sino más bien que la evitemos cuando no *queremos* que se generalice. En los términos del capítulo 4, apelamos a la noción de deseos inconsistentes más que a la de las creencias inconsistentes. Ahora, el argumento de Kant puede interpretarse como diciendo que dado el Dilema del Prisionero (o, para este caso, un Juego de Certeza) no podemos preferir consistentemente el rincón del lado derecho al del lado izquierdo, en la matriz de resultados, ya que esto significaría preferir x a y, y y a x.

14. Esto es contrario del procedimiento de Hughes y Cresswell (1972, págs.77-78) quienes ofrecen una interpretación en la cual un mundo con teléfonos no sería concebible y, por lo tanto, no accesible desde un mundo sin teléfonos.

15. Marx (1867. pág. 163).

16. Véase Schumpeter (1954, pág. 739).

17. Coleman (1973a, pág.67), la bastardilla es nuestra.

18. Vernon (1971, pág. 138).

19. Clark (1899) es usualmente citado como el *locus classicus* para la noción de que cada factor debe ser recompensado de acuerdo con su producto marginal. Con reservas importantes, la teoría de la justicia elaborada por Nozick (1974) puede ser vista como una derivación de esta tradición; véase también la nota 21 de este capítulo.

20. Debe hacerse aquí una breve mención de la llamada "controversia del capital", cuyo resultado más importante es que la utilización del capital homogéneo como un factor de producción es difícil de justificar analíticamente. Véase especialmente Harcourt (1973) y Bliss (1975). Algunos autores, como Nuti (1970), argumentan que la inadecuación de la función $P = f\ (K, L)$ como representación del proceso de producción, implica la necesidad de un funcional $P\ (t) = F(L\ (t))$, en la que el capital ha sido reemplazado por el flujo de inversión de trabajo que lo ha producido. Esto, sin embargo, no puede ser rigurosamente verdadero. Los procesos reales tienen que estar gobernados por causas presentes, solamente. En el nivel ontológico, sabemos que alguna representación tal como $P = f(K_1, K_2 ... K_n, L)$, en la que K_i es capital en algún nivel inferior de agregación, *tiene* que ser posible. (Véase también Elster 1976a). No puede argumentarse desde la *necesidad* del funcional $F\ (L\ (t))$ hacia las conclusiones políticas de la distribución del ingreso, pero la negación de un derecho a la recompensa por el capital presupone la *posibilidad* de este funcional.

21. La noción de principios de resultados finales está tomada de Nozick (1974, págs. 153 y sigs.). Siendo diferentes en todos los demás aspectos, la teoría marxista de la explotación y la de Nozick sobre la justicia comparten la categoría de ser teorías *históricas*, opuestas a las teorías de resultados finales. Como la teoría de la productividad marginal del ingreso es, en sí misma, una teoría de resultado final, la necesidad de las "reservas importantes" en la nota 19 debe quedar clara.

22. Marx tenía un ojo agudo para esto. "*El capitalista individual*...puede elegir la utilización de su capital prestándolo como capital de interés, para producir interés, o la expansión de su valor, utilizándolo como capital productivo, sin importar si existe como moneda, desde el principio, o debe ser convertido a capital moneda. Pero aplicarlo al *capital total* de la sociedad como hacen algunos economistas vulgares, e ir hasta su definición como causa de ganancia es, por consiguiente, absurdo. La idea de convertir todo el capital en capital moneda sin que haya quien compre y utilizar medios de producción que componen el capital total por fuera de una porción relativamente pequeña de éste, existente en dinero, es por consiguiente un sin sentido total."(Marx 1894, págs.373-378, itálicas agregadas).

23. Si el lector se remite al ejemplo de laputanos y ruritanos del capítulo 4 apreciará la relevancia de la siguiente observación:"Los laputanos pueden creer que ellos van a ser atacados por los ruritanos, sin por eso creer que lo creen, pues cada uno de ellos puede ser inconsciente de que los otros lo creen. Nuestra teoría, sin embargo, está diseñada para ser aplicada a lo que un ser humano individual cree (y conoce) y no a lo que se dice que un número de individuos cree (y conoce)" (Hintikka, 1961 pág. 29).

24. Para otras presentaciones históricas de esta idea véase Merton (1936) y Hayek (1967).

25. Desafortunadamente, el trabajo de Haavelmo sobre este tema ha sido fundamentalmente publicado en noruego, con excepción de una breve exposición de 1970. Según parece, buena parte de la teoría sobre bienes públicos y el Dilema del Prisionero fue trabajada por primera vez por Haavelmo, de manera que la historia del pensamiento económico tiene que tener en cuenta este esfuerzo pionero.

26. *Discours sur l'Histoire Universelle* de Bossuet, es de 1681; la primera edición de *Scienza Nuova* de Vico, de 1725.

27. Nozick (1974, pág.19).

28. Keynes (1936, págs. 359 y sigs.) contrasta los siguientes pasajes:"Como esta economía prudente a la que algunas personas llaman *Ahorro* es en las familias privadas el método más certero para aumentar los bienes, imaginan que, así un país sea desértico o productivo, el mismo método, si es seguido de manera general (lo cual ellos piensan que es practicable) tendrá el mismo efecto para una nación, lo cual haría que los ingleses fueran más ricos de lo que son si fueran tan frugales como sus vecinos. Pienso que esto es un error" (Mandeville). "Lo que es prudente en la conducta de toda familia privada, casi nunca puede ser una locura en un gran reino" (Adam Smith). Me parece que el primer pasaje tendría que leerse como una refutación de la falacia de composición en el sentido preciso dado aquí a esa falacia; no es, sin embargo, absolutamente cierto que el segundo pasaje afirme esa falacia *en ese sentido*.

29. Mandeville (1729, vol.I, pág. 89).

30. Mandeville (1729, vol.I, pág. 428).

31. Elster (1975a, págs.193 y sigs.).

32. *Wealth of Nations*, IV:IX

33. Un profundo y muy difícil análisis de la contrafinalidad se encuentra en la *Fenomenología del espíritu* de Hegel, capítulo V.C.a; para interpretaciones (un tanto divergentes) véase Hyppolite (1946, págs.286 y sigs.) y Chamley (1963, págs.23 y sigs.). El concepto de astucia de la razón, por comparación, es esotérico y está más relacionado con la interpretación de la historia de Vico-Bossuet; véase aquí a D'Hondt (1970, págs. 22 y sigs.) y Avineri (1971). Entre los muchos textos, especialmente en la *Ideología alemana* y los *Grundrisse*, en los que Marx expuso la contrafinalidad, la siguiente parece ser una cita particularmente sobresaliente: "Esta dependencia recíproca se expresa en la necesidad constante de intercambio y en el valor de cambio como la mediación general. Los economistas expresan esto de la manera siguiente: cada uno persigue su interés privado y sólo su interés privado; por lo tanto, sirve a los intereses privados de todos, es decir, al interés general, sin desearlo o saberlo. El verdadero punto no es que la búsqueda individual de interés privado promueva el interés de todos, el interés general. Uno bien podría deducir de esa frase abstracta que cada individuo recíprocamente obstaculiza la afirmación de los intereses de los demás, de manera que, en vez de una afirmación general, la guerra de todos contra todos produce una negación general. El

punto es más bien que el interés privado es, en sí mismo, un interés socialmente determinado, que puede lograrse sólo dentro de las condiciones establecidas por la sociedad y con los medios que la sociedad provee. Por lo tanto, está ligado a la reproducción de esas condiciones y medios. Es el interés de personas privadas; pero su contenido, tanto como la forma y los medios para realizarlo, está dado por condiciones sociales independientes de todos" (Marx, 1857-58, pág.156). Obsérvese que en este pasaje Marx propone dos puntos conceptuales importantes: primero, hay una diferenciación entre la contrafinalidad y la mano invisible; segundo, la idea de que los "intereses privados" no se forman al azar, opuesto a lo que sugiere Parsons (1949, pág.493), quien habla acerca de "la presunción implícita del azar de los fines últimos"en el pensamiento de Marx.

34. *Wealth of Nations*, IV:IX.

35. Forrester (1969)

36. Schelling (1971, pág.150).

37. Sweezy (1962).

38. Sartre (1960, págs.232 y sigs.).

39. Véase por ejemplo Benedict (1972, págs. 81 y sigs.) y Neher (1971), respectivamente, para evidencia empírica y un análisis teórico.

40. Finley (1973a. pág.175).

41. Genovese (1965, pág.165). Se muestra en el capítulo 6 que, como afirmación empírica, no puede ser sustanciada. Parecería que Genovese ha tenido mejores intuiciones acerca de las contradicciones mentales generadas por la esclavitud que sobre las contradicciones sociales.

42. Párrafos 243-246. Aquí me apoyo en la interpretación ofrecida por Weil (1966, pág.99).

43. Véase Simon (1954a) y Brams (1976, cap.3) sobre las discusiones de profecías autocumplidas. En estas discusiones se supone que la votación real **y** es una función $f(\mathbf{x})$ de la votación predicha **x** en la que **x** e **y** son vectores que indican los porcentajes de votos que van a los diversos partidos. El conjunto de dichos vectores es compacto y convexo, y si la función f es continua, podemos invocar el teorema del punto fijo de Brouwer, para afirmar la existencia de una predicción $\overline{\mathbf{x}}$ tal que $\overline{\mathbf{x}} = f(\overline{\mathbf{x}})$, es decir, una predicción que se toma en cuenta a sí misma. Ahora hay dos problemas inherentes a esta elegante y justamente famosa demostración. En primer lugar, como ya se observó en el capítulo 3, la función $f(\mathbf{x})$ tiene que hacerse accesible al público, pues si a los votantes se les dijera cómo iban a reaccionar a la predicción, su reacción y también el punto fijo (en el caso típico) cambiaría. De manera que el mismo problema surgiría de inmediato. Para que el método funcione, *algo tiene que ser guardado en secreto*, así sea la predicción de primer orden de votar, la predicción de segundo orden de cómo la predicción de primer orden va a influenciar la votación real, la predicción de tercer orden de cómo la predicción de segundo orden va a influenciar la reacción a la predicción de primer orden, etcétera. En segundo lugar, la presunción de continuidad es ciertamente debatible. Para un ejemplo que resalte esto, podemos tomar el caso en que la encuesta de opinión da una distribución continua sobre algún tema, más bien que una repartición discreta de votantes en los partidos. Un reciente trabajo sobre la teoría de la catástrofe (Isnard y Zeeman, 1974) ha demostrado cómo pueden surgir discontinuidades en este caso. Permítasenos suponer que la distribución de la opinión es inicialmente bimodal y que los individuos reaccionan a la distribución publicada acercándose más al modo más cercano a su opinión. Si la distribución cambia continuamente por una unimodal, habrá discontinuidad (una catástrofe) en la función de reacción en el punto donde uno de los modos desaparece. Exactamente el mismo tipo de discontinuidad puede surgir en el caso vector-espacio.

44. Para la curva de oferta de tendencia regresiva véase Chayanov (1966, págs. 80 y sigs.).

45. Un sistema dinámico es estructuralmente estable si el comportamiento global del sistema permanece igual ante cualquier perturbación lo suficientemente pequeña de los parámetros del sistema; en el caso de $b=d$ es claro que el más pequeño cambio en b o d cambiará las propiedades globales de las trayectorias.

46. Allen (1966, págs. 13-14).

47. Allen (1966, pág. 197).

48. Coddington (1968, págs.58 y sigs.); véase también página 64 para un comentario explícito acerca de la diferenciación que he intentado expresar contrastando la primera y tercera tesis sobre las contradicciones en el capítulo anterior. Véase también a Leijonhufvud (1966, pág.103) para una observación similar.

49. Marx creía que la producción industrial se había llevado a cabo con coeficientes fijos, de manera que la sustitución de factores no era una posibilidad relevante. La adjudicación a Marx de una teoría de coeficientes fijos es cuestionada algunas veces (véase, por ejemplo, Samuelson, 1957) pero la documentación aducida por Maarek (1975, págs.42-43), a las cuales podrían agregarse otros pasajes suficientes para eliminar toda duda.

50. Esto fue primeramente reconocido, según parece, por Salter (1960, págs.43-44). Un reconocimiento parcial del mismo hecho se encuentra en Sweezy (1962, pág.88) quien ve que el efecto reductor de salarios de la mecanización es un punto extra o un efecto colateral de las acciones llevadas a cabo por otros motivos, pero no explica cuáles pueden ser esos otros motivos, por qué las innovaciones tendrían una tendencia a ahorrar trabajo.

51. Marx (1867, pág.435); Smelser (1959, pág.235).

52. Fellner (1961, pág.307).

53. Véase la discusión de aprendizaje "irracional" versus "racional" en las observaciones del modelo de la telaraña.

54. David (1975, págs.36 n.2 y 54). David es consciente del problema del que actúa por sí solo, que surge tan pronto como cada capitalista comprende que los demás capitalistas están actuando de manera de lograr beneficios sin importar lo que él haga, pero esta discusión no enfatiza que el juego emprendido no tiene solución.

55. Véase Morishima (1973), para este problema.

56. Véase, por ejemplo, Takayama (1974, capítulo 4), o Von Weiszäcker (1971b).

57. Creo que se pueden distinguir entre cuatro distintos modos de la elusiva idea de la "teoría del valor del trabajo":

(a) Los precios son proporcionales a los valores.

(b) La suma de todos los precios es igual a la suma de todos los valores.

(c) Los valores pueden ser determinados independientemente de los precios, pero no viceversa.

(d) Los precios son independientes de la composición de la demanda final.

La identidad local de los precios y los valores implícita en la interpretación (a) es afirmada en *El capital* I para ser abandonada en *El capital* III. Como bien se sabe actualmente, esta no es la "gran contradicción" que Böhm-Bawerk (1898) pensaron, sino sólo un caso de análisis por aproximaciones sucesivas. La identidad global de precios y valores es afirmada en *El capital* III, pero como los valores son absolutos y los precios, relativos, esto es o bien relativamente insignificante, o bien sólo la definición de un numéraire. La interpretación (c) surge de la diferenciación hegeliana entre la esencia (valores) y la apariencia (precios), mientras que la interpretación (d) vuelve al ancestro ricardiano del marxismo. Para un análisis preciso de las condiciones bajo las cuales las interpretaciones (c) y (d) son válidas véase Hoel (1974).

58. "Lo que determina el valor no es la cantidad temporal de trabajo incorporado a los productos sino más bien la cantidad temporal de trabajo necesario, en un momento dado" (Marx,1857-58, pág.135). Para comentarios acerca del problema general de las teorías del valor simultáneas versus las históricas, véase Dorfman y cols.(1958, pág.234); para la utilización por Marx del "razonamiento circular" para defenderse de sus críticos, véase Bródy (1974, pág. 84).

59. Para esta "subsunción real del trabajo bajo el capital", véase especialmente Marx (1933).

60. El primer rótulo es utilizado por Olson (1963, pág.131), el segundo por Barry (1973, pág.118).

61. Hobsbawm (1954, pág. 56). El mismo punto está abundantemente elaborado en North y Thomas (1973, *passim*).

62. Thompson (1971, pág. 120).

63. Williams (1966, pág. 209)

64. Salthe (1972), pág. 339)

65. Williams (1966, págs. 212 y sigs.). En cualquier cardumen de una densidad dada, lo mejor para cualquier pez es estar en la mitad del banco donde está más protegido de los predadores, pero a mayor número de peces que busquen ir al medio, mayor la densidad del cardumen y mayor la vulnerabilidad a los predadores.

66. Merton (1957), Stinchcombe (1968).

67. Véase Glantz y Albers (1974) para un estudi de caso.

68. Mandeville (1729, *passim*) Marx (1862-63, vol. 3, págs. 363-364).

69. Véase Winter (1964) y Winter (1975) para exploraciones detalladas de este modelo.

70. Me parece que la mayor parte de los lectores se han encontrado con un número más que suficiente de ejemplos de este razonamiento, así que la documentación es superflua. Podemos observar que el mismo Marx (véase, por ejemplo, Marx 1871, págs. 338, 541, 592, 594) utilizó esta línea de argumentación en su análisis del Segundo Imperio en Francia; en mi opinión no es una de sus glorias. De hecho, el análisis empírico de Marx acerca del Estado capitalista tiende a oscilar entre el humor conspirativo, como en las diatribas contra Palmerston en Marx (1857, págs. 278 y sigs.) y el vacuamente funcional.

71. Véase Elster (1977a) para una mayor discusión.

72. Popper (1963, pág. 249).

73. Kenny (1970) hace una distinción entre previsión e intención, pero creo que, en el presente caso, las dos nociones coinciden. La razón por la que podemos pensar que los jugadores en un Dilema del Prisionero no intentan realmente (o sea, desean) un resultado subóptimo, es porque tendemos a comparar las entradas diagonales en la matriz de resultados, pero desde el punto de vista de los actores son relevantes sólo las comparaciones horizontales y verticales.

74. Luce y Raiffa (1957, pág. 97).

75. Véase Olson (1963); Riker y Ordeshook (1973, cap. 9); Staaf y Tannian (1972); Arrow y Scitovsky (1969).

76. Por lo tanto, soy bastante escéptico acerca de la siguiente afirmación: "Como ningún señor [feudal] logró captar siquiera una pequeña porción del producto social del proceso agrícola mejorado y tecnificado, cada uno tenía poco incentivo para intentarlo. La respuesta racional sería esperar a que otro terrateniente sobrellevara los costos de la investigación y el desarrollo y luego simplemente imitar su procedimiento, si hubiera tenido éxito. Cada señor, por consiguiente, eligió esperar y se logró progresar poco" (North y Thomas, 1973, pág. 62).

77. Boudon (1974, pág. 198).

78. Boudon (1977, cap. IV).

79. Boudon (1977, pág. 265).

80. Véase Elster (1976b).

81. Boudon (1977, cap.V).

82. Véase Stonffer y cols.(1949) y las explicaciones ofrecidas por Merton (1957, págs. 225 y sigs.) y Hirschman (1973, pág.552, n. 9).

83. Phelps y Pollak (1968), sobre lo que me apoyo fuertemente para lo que sigue. Véase también Elster (1976d).

84. Sejersted (1975); véase también Elster (1976d).

85. Phelps y Pollak (1968). Su resultado es realmente bastante ambiguo y el lector interesado debe consultar el original para mayores detalles.

86. Este es un caso de la "explotación de los grandes por los pequeños", como es discutido por Olson (1963, pág. 29).

87. Gallagher y Robinson (1953). En este artículo y en la discusión siguiente, el término ha sido utilizado para denotar el imperio informal que *acompañó* al librecambio, mientras que yo lo utilizo aquí para el control informal obtenido *por medio* del librecambio. El término es también utilizado en este último sentido por Kindleberger (1975, pág.33 y sigs.), refiriéndose y citando a "los economistas políticos que buscaron el librecambio como medio para disminuir el ritmo de desarrollo de la manufactura en el continente". No es necesario enfatizar la analogía con la inversión del librecambio imperialista ejercida hoy por los países productores de

petróleo, en el que los países industrializados importadores de petróleo han encontrado de su interés imponer un precio *mínimo* de compra del petróleo para hacer posible el desarrollo de fuentes alternativas de energía. A un contrincante que intente arruinar al otro vendiendo barato, es racional replicarle comprando caro.

88. Schumpeter (1953, pág. 83) expone el problema en su generalidad total: "Un sistema —cualquier sistema, económico u otro— que en *cualquier* punto dado del tiempo utiliza plenamente sus posibilidades para obtener mayor beneficio, a la larga puede ser inferior al sistema que lo hace en un punto *no* dado temporal, porque si este último fracasa, esto puede ser una condición para el nivel o la velocidad de ejecución a largo plazo." Una elaboración de la idea es ofrecida por Arrow (1971a). El argumento ha sido extendido hacia el comportamiento político, en el que hay toda una escuela de pensamiento que, por ejemplo, ve la apatía (suboptimalidad política de corto plazo) como una condición para la democracia (suboptimalidad política de largo plazo). Una muestra de esta posición y del nivel de crítica levantado contra ella puede verse en las dos siguientes citas: "El descubrimiento de que los ciudadanos normalmente no utilizan más que una fracción de sus recursos políticos fue originalmente una sorpresa y una decepción para los científicos políticos, a quienes se les había llevado a creer que la democracia requiere para su funcionamiento la mayor participación posible de todos sus ciudadanos. Pero, pronto se encontró que cierto grado de apatía tenía algunas ventajas compensatorias, en la medida en que contribuye a la estabilidad y flexibilidad de un sistema político." (Hirschman, 1970, pág. 14). "Mi argumento es que, lejos de ser una condición saludable y necesaria de la democracia, la apatía es una respuesta de aislamiento frente al desequilibrio en el acceso de los diferentes grupos de interés a quienes toman las decisiones." (Finley, 1973b, pág. 67). Simpatizo aquí con Finley, aunque pienso que hay una confusión en el pasaje citado entre las causas y los efectos de la apatía; pero la misma confusión existe en la raíz del análisis funcional, que ve los efectos estabilizadores de la apatía como la causa de su persistencia. (No sé si Hirschman aceptaría este análisis.) A no ser que se piense de mi crítica al radicalismo y la impaciencia política en la última sección del capítulo que también hace de la apatía una virtud, quisiera señalar la diferencia entre la restricción autoimpuesta y la apatía inducida externamente.

89. La importancia del sistema de patentes ha sido enfatizada recientemente por North y Thomas (1973) quizá como *la* principal condición para el surgimiento del capitalismo y de la Revolución Industrial. En mi trabajo sobre Leibniz (Elster, 1975a, capítulo III) discuto con cierto detalle cómo se articulaba en Alemania, a finales del siglo XVII, el dilema entre las suboptimalidades estáticas y dinámicas. La industria del conocimiento en ese tiempo y lugar era subóptimo desde ambos puntos de vista: los artesanos guardaban sus secretos tan celosamente que la utilización estática era peor de lo que hubiera sido con el sistema de patentes, y al mismo tiempo, la ausencia de un sistema de patentes dio pocos resultados al conocimiento y un débil incentivo para innovar. Leibniz propuso un sistema por el cual los artesanos podían ser forzados a revelar sus secretos por una compensación del Estado. Un punto crucial sería el de los principios que gobernaran el monto de la compensación. Un sistema ideal sería ligar la recompensa a los aumentos en la productividad que el invento trajera consigo, pero Leibniz parecía tener en mente un proyecto por el cual el artesano recibiría una cantidad equivalente a lo que hubiera obtenido si hubiera guardado el conocimiento de su habilidad para sí. Este último sistema eliminaba la suboptimalidad estática y no producía mayor diferencia desde el punto de vista dinámico.

90. Véase Arrow (1971a, pág. 147 y *passim*). La desviación de la maximización de la utilidad a partir de la maximización de la ganancia puede ocurrir incluso en ausencia de riesgo, por ejemplo, si el intercambio del ingreso por ocio del empresario cambia con el nivel de ingresos; véase, Quirk y Saposnik (1968, pág. 35). Como observó Scitovsky (1971, pág. 153) esta fuente de desviación de la utilidad de la ganancia es menos importante en el caso del empresario "puritano" clásico que en el tipo del gerente moderno, mientras que lo opuesto es presumiblemente verdadero para la fuente mencionada en el texto.

91. Young (1976) ofrece un análisis extenso de esta frase, el cual, si es valioso en el nivel exegético, no es, para mi criterio, muy convincente en el nivel conceptual. Niega que la noción marxista de una contradicción tenga que ver con la contradicción lógica y argumenta que

tendría más bien que ser entendida como "la relación entre dos elementos de una estructura polar o dos fases de un proceso polar". Desafortunadamente, esta noción no se clarifica con los ejemplos aportados y no creo que logre el nivel de un concepto intersubjetivo u operacional.

92. Estas cuestiones exegéticas o terminológicas se discuten más aún en Elster (1972).

93. Marx (1867, pág. 393).

94. Arrow (1971a, págs. 156 y sigs.).

95. "La rapidez del cambio en la composición orgánica del capital y en su forma técnica aumenta" (Marx, 1867, pág. 631).

96. Para la comprobación de esta interpretación, señalaría especialmente el siguiente pasaje: "Aunque limitado por su propia naturaleza [el capital] se empeña en el desarrollo universal de las fuerzas productivas y por lo tanto se transforma en la presuposición de un nuevo modo de producción, que está fundado, no en el desarrollo de las fuerzas productivas, con el fin de reproducir o por lo menos expandir una condición dada, sino donde el libre, sin *obstrucción*, progresivo y universal desarrollo de las fuerzas productivas sea él mismo la presuposición de la sociedad y por lo tanto de su reproducción" (Marx 1857-58, pág. 540). La palabra que he puesto en itálicas muestra cómo Marx no previó ningún problema de carencia de incentivos para innovar en la sociedad comunista. Lo que haría necesaria una explicación es la baja tasa de actividad innovadora en las sociedades precomunistas más bien que la alta tasa que se generalizará en el comunismo.

97. Gerschenkron (1966, pág. 8).

98. Lancaster (1973).

99. Estableciendo x_1 para el ritmo de consumo del producto por los trabajadores, x_2 para la tasa de reinversión de las ganancias y u_1 para el límite superior de x_1, podemos diferenciar entre los siguientes modos de explotación:

(1) u_1 bajo, x_2 bajo.
(2) u_1 bajo, x_2 alto.
(3) u_1 alto, x_1 bajo, x_2 bajo.
(4) u_1 alto, x_1 bajo, x_2 alto.

La primera forma es típica de las sociedades precapitalistas (si generalizamos el método e incluimos la explotación en las sociedades esclavistas y el feudalismo), la segunda es típica del capitalismo en sus inicios, la última del capitalismo moderno. La tercera no es muy factible, porque los trabajadores pueden vengarse con x_1 alto contra x_2 bajo.

100. Si la no provisión de bienes públicos pudiera ser una forma de contrafinalidad más bien que de suboptimalidad, podríamos tener un contraejemplo de esta proposición, pero he dado mis motivos para dudar de esta posibilidad. De nuevo reitero que no hay nada sustancial en juego aquí.

101. Bateson (1972, pág. 305)

102. Owen (1968, pág. 72)

103. Luce y Raiffa (1957, pág. 107)

104. Rémond de Montmort escribió en su clásico tratado *Essai d'Analyse sur les Jeux de Hasard* (2a. edición, 1713): "Estas cuestiones son muy simples, pero creo que no pueden resolverse. Si se prueba que éste es el caso, será una gran pena, ya que esta dificultad aparece muy frecuentemente en la vida civil. Por ejemplo, cuando cada una de dos personas en alguna transacción (*ayant affaire ensemble*) quiere alinearse con el comportamiento del otro." Más recientemente, Latsis (1976, pág. 31) escribe:"En el oligopolio, cada uno de los que toman dcisiones tiene que adivinar el comportamiento esperado del otro; y si se supone conocimiento perfecto de ambos lados, esto llevaría a una regresión infinita." (Siendo justos con el autor, debería mencionarse que el juego del oligopolio es, de hecho, un juego sin solución, pero ni la oración citada ni su contexto muestran ninguna conciencia de la diferenciación entre los juegos, con o sin solución no cooperativa.) Dichas exposiciones obvian tanto la posibilidad de llegar a una solución, incluso cuando no surge ninguna estrategia dominante, como la noción de estrategias mixtas.

105. Arrow (1971a, pág.184).

106. Hirschman (1970, pág.18), citado de Hirschman (1958, pág 63).

107. North y Davis (1971, pág.10).

108. Para la noción de ultraestabilidad véase Ashby (1960).

109. Adherentes muy conocidos de este punto de vista son Easton (1965) y Deutsch (1963). Para una utilización explícita de la noción de ultraestabilidad, véase Kaplan (1967, pág.161): "El sistema político tiene la capacidad de metatarea de actuar como regulador ultraestable de todo el sistema en el cual funciona." La falacia básica del método consiste en creer que las sociedades son sistemas. Podemos hablar de manera útil acerca de los sistemas que han sido diseñados teleológicamente por el hombre o teleonómicamente por selección natural, pero las sociedades no son el resultado del diseño ni de la evolución que, en algún sentido, las haga asemejarse a la evolución natural. Los criterios de identidad de las sociedades están mal definidos, al igual que los criterios de "salud", "enfermedad", "nacimiento" o "muerte" de las sociedades. Si queremos, podemos hablar acerca de las sociedades que se adaptan a los problemas, pero esta adaptación difiere de la adaptación orgánica (i) en que no es automática y (ii) en que no está restringida a un conjunto de respuestas que han sido programadas por la evolución.

110. La percepción y organización, por ejemplo, son discutidas bajo el mismo encabezamiento por North y Davis (1971, pág.57).

111. Marx (1867, pág.239). El único autor (de acuerdo con mi conocimiento) que ha llamado la atención sobre las implicaciones curiosamente no marxistas de este pasaje es Smelser (1979, pág.393).

112. "Mientras todo vaya bien, la competencia efectúa una fraternidad operativa de la clase capitalista, como hemos visto en el caso de la ecualización de la tasa general de ganancia, de manera que cada uno comparte en la rapiña común una porción proporcional al tamaño de sus respectivas inversiones. Pero, tan pronto como deja de ser una cuestión de compartir ganancias para compartir pérdidas, cada uno intenta reducir su propia participación a un mínimo, para arrojárselas a otro. La clase, como tal, tiene que perder inevitablemente. Qué parte de la pérdida debe sobrellevar el capitalista individual, es decir, hasta qué punto tiene que compartirlo todo, se decide por la fuerza y la astucia, y la competencia se transforma, entonces, en una pelea entre hermanos hostiles. El antagonismo entre los intereses de cada individuo capitalista y los de la clase capitalista en conjunto asoma en la superficie" (Marx, 1894, pág.253). Este pasaje es bastante confuso, ya que la última frase sugiere un juego de suma variable entre los capitalistas, y las afirmaciones precedentes, un juego de suma cero. Sobre la base de otros pasajes, creo que debe enfatizarse la última frase.

113. Marx (1867, pág.763).

114. Marx (1867, pág.506).

115. Marx (1867, pág.462).

116. Thompson (1968, págs.566-567); Perkins (1969, pág.188).

117. Genovese (1965, pág.176).

118. Gerschenkron (1965, pág.710); Portal (1965, pág. 804).

119. Smith (1968)

120. Selden (1971, págs.35 y sigs. y págs.54 y sigs.)

121. Simmel (1908, págs.460 y siguientes).

122. Olson (1963, pág.28).

123. Granovetter (1973, págs.1373 y sigs).

124. Una breve revisión de los trabajos producidos por una investigación de la literatura puede ser útil para los lectores. Para la primera edición de *Class, Status and Power*, Seymour Lipset y Joan Gordon (1953) escribieron un trabajo que toca varios aspectos de la relación entre el promedio de cambio de pertenencia y la acción colectiva. El artículo no se mantuvo en la segunda edición, quizá porque algunas observaciones sobre la relación de la movilidad y la ausencia de conciencia de clase no fueron confirmadas en el último trabajo de Lipset y Bendix (1959). Otro trabajo inicial importante es el de Kerr y Siegel (1964), quienes encontraron que la movilidad horizontal era un factor clave en el analisis de la propensión a la huelga. Más recientemente, Thernstrom (1970) argumenta sobre la base de lineamientos similares. Fellin y Litwak (1963) estudiaron el impacto de la movilidad sobre la cohesión de un vecindario, pero la capacidad para la acción colectiva entra sólo marginalmente en la última

noción. Un conocido trabajo de Germani (1966) sobre diversas consecuencias de la movilidad tiene, sorpresivamente, poco que decir acerca de nuestro problema; esto es válido, también, para el trabajo de North y Davis (1971) sobre las condiciones para el cambio institucional. El trabajo de Granovetter (1973) es muy estimulante, pero utiliza un método un tanto indirecto para este trabajo.

125. Olson (1963, pág.69).

126. Lipset y Gordon (1953).

127. Wilensky y Edwards (1959) discuten "el deslizador" y sus acomodamientos ideológicos con algún detalle.

128. Simmel (1908, pág. 431). El pasaje citado es una traducción más o menos libre de lo siguiente: *Der prinzipielle Demokrat aber will nicht beherrscht werden, selbst um den Preis, dass ihm damit gedient wird.* Véase también Peyrefitte (1976, pág.42).

129. Lipset y Gordon (1953).

130. Arrow (1971b, pág.22).

131. Finley (1973a, pág.77).

132. Bernstein (1899, págs. 73 y sigs).

133. Luxemburg (1899, págs.386-387).

134. Thalheimer (1930).

135. Thompson (1971, pág.135).

136. Bjørgum (1976) aporta evidencia para Noruega. Kerr y Siegel (1964, n.10) citan los "itinerantes, pero ocupacionalmente especializados, esquiladores de ovejas en el oeste de los Estados Unidos", como otro ejemplo de un grupo en el que la movilidad horizontal no impide el desarrollo de una alta conciencia de clase.

137. Véase Meisner (1967, pág. 169) y Elster (1977) para otras observaciones sobre este punto.

138. Me refiero al trabajo de Sivert Langholm y sus discípulos, especialmente Jan Myhre.

139. Thernstrom (1970, pág.227).

140. Si adoptamos el modelo de movilidad de Cornell (para esto véase Boudon,1973, págs. 85-95), el desempleo tiende a crear las condiciones para su propia perpetuación ("inercia acumulativa") en el sentido de que mientras más tiempo haya estado una persona sin trabajo, menores se hacen las posibilidades de que consiga un trabajo en el siguiente período. (Suponemos que éste es un efecto posterior real y no sólo un efecto de muestreo.) En la medida en que esto se debe a los cambios de la persona misma (y no a cambios de actitud en los empleadores potenciales), presumiblemente tomará la forma de desidia creciente y falta de motivación, lo cual, de nuevo, significa que en el grupo de desempleados, las personas con mayor tiempo de pertenencia no necesariamente van a estar motivadas para la acción colectiva.

141. Thernstrom (1970, págs.234-235) (cita corriente).

142. Marx (1894, pág.601).

143. Girod (1971, págs.51 y sigs. y *passim*). Véase también Thernstrom (1973, págs. 77, 94 y sigs.) y Blau y Duncan (1967, págs.54 y sigs).

144. Mukherjee (1954, pág.286). Para otras observaciones acerca de la movilidad en más de dos generaciones véase Schumpeter (1951, pág.169) y Lipset y Bendix (1959, págs. 74-75).

145. Brooke (1972, pág.93).

146. Bachrach y Baratz (1963).

147. Granovetter (1973).

148. North y Thomas (1971, pág.60).

149. Véase Luce y Raiffa (1957, págs.124 y sigs).

150. Arrow (1974, pág.42).

151. Para posteriores lecturas están recomendados los trabajos corrientes de Luce y Raiffa, 1957 y —en una vena más matemática—Owen, 1968.

152. Véase especialmente Sen, 1967, para extensiones a *n* personas del Dilema del

Prisionero y del Juego de Certeza.

153. Las observaciones anteriores se apoyan fuertemente en Luce y Raiffa (1957, págs. 106 y sigs). Sin embargo, creo que el concepto de solución ofrecido aquí es más natural que la "solución en el sentido estricto" y la "solución en el sentido de Nash", que son las principales alternativas. El Juego de Certeza no se puede resolver en el sentido de Nash, mientras que el Dilema del Prisionero no se puede resolver en el sentido estricto. Sin embargo, sobre bases intuitivas, creo que ambos juegos pueden resolverse, en el sentido de que los actores racionales plenamente informados convergerán hacia un modo de comportamiento único.

154. Para el Dilema del Prisionero véase Rapoport y Chammah (1965) y Taylor (1976). Para el Juego de Certeza véase Sen (1967) y Sen (1973). Para la Batalla de los Sexos véase Luce y Raiffa (1957, págs. 90 y sigs., págs.115 y sigs.). Para el Juego de la Gallina véase Rapoport (1966, págs. 137 y siguientes).

155. Para una discusión de una peculiar "cuasi solución" en la Batalla de los Sexos y en el Juego de la Gallina véase Luce y Raiffa (1957, págs. 93-94) y Rapoport (1956, págs. 140 y sigs).

156. Véase Luce y Raiffa (1957, capítulo 2) para una discusión de la teoría de la utilidad, como es empleada en la teoría de los juegos.

157. Véase Rapoport y Chammah (1965 y col.) y Rapoport (1976), para algo de su trabajo acerca de la teoría experimental del juego.

158. Rapoport y Chammah (1965, págs.123 y sigs).

159. Howards (1971) criticado (correctamente, en mi opinión) por Taylor (1976, págs.64 y sigs).

160. Shubik (1970), Taylor (1976); véase también Taylor (1977) para extensiones a la biología.

161. El argumento para esta conclusión es muy simple. Cada jugador sabe que en el último juego ambos escogerán la estrategia no cooperativa, porque después de ese juego no hay futuro en el cual la cooperación pueda ser recompensada. Eso, sin embargo, significa que el penúltimo juego efectivamente se vuelve el último y la misma consideración se aplica aquí y también hacia atrás en el tiempo, hasta el primer juego. Creo que esta noción puede ser generalizada: ¿no daríamos algunos años de nuestra vida por mantenernos ignorantes del número de años que tenemos para vivir?

162. Luce y Raiffa (1957, pág.102).

163. Taylor (1976, pág.93).

164. Citado por Luce y Raiffa (1957, pág. 165).

165. Luce y Raiffa (1957, pág.165).

166. Las mutaciones, la senectud y el cáncer son fenómenos subfuncionales (Williams, 1966, págs.138 y sigs, 225 y sigs); los fenómenos suprafuncionales fueron discutidos en los ejemplos biológicos de la falacia estructuralista.

167. Davidson (1973).

168. Elster (1977). Se mencionan aquí dos características que separan la adaptación intencional de la funcional: la capacidad humana para obtener lo máximo global, aunque esto requiera esperar o, incluso, abstenerse temporalmente un maximo local, y la capacidad humana para alcanzar la solución en un juego en el que no hay estrategias dominantes.

169. Por lo tanto, estoy ampliamente de acuerdo con la siguiente observación: "La fastidiosa debilidad del sistema marxista, omisiones aparte, es para este escritor lo que muchos otros críticos consideran una base de fuerza, esto es, su estructuralismo, o sea su tendencia a tomar las relaciones importantes como técnicamente determinadas detrás de los sistemas de precios, dejando a estos últimos con pocas funciones, más allá de igualar las ganancias" (Bronfenbrenner, 1965, pág. 223). Es justo, sin embargo, añadir que la teoría marxista de la tasa decreciente de ganancia es—dejando de lado los errores matemáticos— un modelo de análisis centrado en los actores.

170. Becker (1962).

171. Cobb (1973). Tullock (1974).

172. Tullock (1974).

173. Innumerables discusiones recientes, con frecuencia sutiles y fascinantes y algunas veces concebiblemente válidas, podrían citarse para probar la justificación racional subyacente en el comportamiento aparentemente irracional. En verdad creo que como una idea reguladora en las ciencias sociales, uno debería siempre comenzar por buscar algún *significado* o *fin* detrás de cada comportamiento, pero uno tendría que esforzarse para *examinar* las explicaciones racionales de elección, confrontándolas con otras explicaciones, en vez de postular sólo entidades inobservables (tales como "costos psíquicos"), que hacen que todo el comportamiento resulte racional por definición. En particular, si se habla de una *elección*, esto es en sí mismo un acto que ocurre en el tiempo y el espacio y no algo que pueda ser adjudicado a sus resultados observados sobre la presunción de racionalidad. Para dar un ejemplo de un análisis que, de acuerdo con mis conocimientos, es definitivamente equivocado (o vacío) se puede citar lo siguiente: "Las personas actúan en términos de posiciones ideológicas generales más que incurrir en el costo de adquirir información sobre un tema particular, incluso cuando para ellos sería rentable hacerlo. Por lo tanto, reaccionan como liberales, conservadores, radicales o reaccionarios. Las ideologías son una forma de economizar en el costo de la información y, por lo tanto, son una respuesta racional a los costos de información sobre el amplio espectro de temas que enfrentan" (North, 1971, pág. 122). En primer lugar, en este contexto, ¿"respuesta" significa un hecho fechable o es solamente una invención analítica? En segundo lugar, las personas *¿nunca* actúan estúpidamente como producto de la estupidez y rígidamente como producto de la rigidez? En tercer lugar, ¿cómo obtienen la información que les posibilita economizar óptimamente los costos de información?

174. Una aproximación interesante, que sólo se relaciona de manera indirecta con el presente problema, es la de McFarland (1976). Un enfoque más directamente orientado hacia nuestro problema es el de Coleman (1973b) quien se equivoca, sin embargo, confundiendo la intencionalidad sustantiva de las acciones humanas con la "intencionalidad como—si" de los procesos físicos, tal como la intencionalidad aparente exhibida por el principio del menor tiempo en la física.

175. Simon (1954b).

176. Véase, por ejemplo, Riker y Ordeshook (1973, capítulo 2).

177. Winter (1964, pág.262).

178. Winter (1965, pág.83).

179. *Etica nicomaquea*, 1113-1114.

180. Elster (1975a pág.72 n.85). El problema es el siguiente. Por una parte, Descartes dice que Dios no está limitado —en su elección del mundo— por ningún criterio de perfección preexistente; pero por otra parte, Descartes también intenta justificar el mal del mundo con los argumentos tradicionales. La contradicción puede ser resuelta suponiendo que el mundo actual es el mejor de los mundos posibles, *de acuerdo con sus propios (creados) criterios de perfección.*

6

Los contrafácticos y la nueva historia económica

La noción de proposiciones contrafácticas (o contrarias a los hechos) recientemente se ha hecho importante en dos disciplinas que no están relacionadas: la lógica filosófica y la historia económica. En ambos casos, el interés en los contrafácticos surge de una preocupación por la causación: en el caso del filósofo, por el *concepto* de causalidad, y en el caso del historiador, por la *evaluación* de la importancia relativa de las causas. Explicaré brevemente cómo estas preocupaciones han llevado a la noción de los contrafácticos. Luego proseguiré con algunos aspectos especiales de la noción, ofreciendo una historia breve y esquemática del concepto, un análisis tentativo de los contrafácticos históricos y, finalmente, un número de ejemplos extraídos de recientes controversias en historia económica.

El punto de partida para todos los análisis de la causación es el punto de vista de Hume, de que ésta significa una conjunción constante entre eventos. Desde entonces, esta "teoría de la regularidad de la causación" ha sido redefinida de diferentes maneras, de tal forma que "en los análisis de regularidad del presente, una causa es definida (a grandes trazos) como cualquier miembro de cualquier conjunto mínimo de condiciones actuales que son conjuntamente suficientes, dadas las leyes, para la existencia del efecto".[1] Ahora bien, en la teoría de Hume existe una dificultad obvia para distinguir la causación de la mera correlación, y surgen problemas similares en el caso de las versiones más elaboradas.[2] Para distinguir la causación de la correlación podemos señalar que la primera requiere que se especifique que, si la causa no hubiera ocurrido, el efecto no habría ocurrido, mientras que en el caso de la segunda, no hay implícito un contrafáctico de ese tipo. A la mayoría de los filósofos, *definir* la causación por medio del concepto de los contrafácticos les ha parecido un caso de *obscurium per obscurius*, porque el estatus lógico de los contrafácticos no ha sido bien entendido. En particular, varios de los problemas que le crean dificultades

a la teoría de la regularidad de la causación parecen también constituir obstáculos insuperables para la mayoría de las explicaciones tradicionales de los contrafácticos. A partir del surgimiento reciente de la teoría de los contrafácticos en términos de mundos posibles, este enfoque de la causación ha comenzado a parecer plausible. Los detalles de este enfoque serán especificados a continuación.

La razón por la cual al historiador le interesan los contrafácticos es bastante diferente. Para empezar, observemos que tradicionalmente el historiador ha sostenido dos creencias incompatibles. Por un lado, ha asumido como un hecho que algunas causas son más importantes que otras, pero por el otro, ha estado acostumbrado a hablar de "la tela sin costuras de la historia" y a criticar a los pocos historiadores que se recrean especulando sobre "lo que podría haber ocurrido si". Ahora bien, estos dos principios son difícilmente reconciliables, ya que en una disciplina no experimental y no comparativa uno difícilmente pueda discutir la importancia relativa de las causas sin meterse en algún tipo de experimento intelectual donde quite sucesiva y separadamente cada una de las causas en cuestión y evalúe qué diferencias habría ocasionado la ausencia de esta causa al fenómeno estudiado.[3] Algunos historiadores, por lo tanto, han llegado a admitir que habían estado hablando de manera contrafáctica todo el tiempo sin reconocerlo, pero no sé si ya son una mayoría en la profesión. La controversia alrededor de los contrafácticos en la historia no depende de la elección que hagamos entre las teorías de la causación —discutidas en el párrafo anterior—. Si aceptamos una teoría contrafáctica de la causación —sea en una de sus variedades tradicionales o en la variedad de los mundos posibles— por supuesto estamos obligados a decir que la causación incluye a los contrafácticos, pero esta obligación también se aplica a los que proponen la teoría de la regularidad. Con respecto a la última teoría, cualquier declaración sobre causalidad *implica* una declaración contrafáctica, aun cuando la teoría no acepta que la causación debería ser *definida* en términos contrafácticos. Lo que ha preocupado a los historiadores, la idea de que los enunciados sobre causación implican declaraciones sobre contrafácticos, ha sido dada por descontada por los que trabajan en lógica, y el problema de estos últimos, si la causación debiera ser definida en términos de contrafácticos o si el vínculo es solamente no definicional, ha sido demasiado esotérico para los primeros.

Tres capítulos sobre la historia del concepto

El problema lógico de los contrafácticos y los mundos posibles tiene un trasfondo histórico complejo. Aun si tuviera el espacio, me falta la capacidad para ofrecerle al lector una versión adecuada. Por lo tanto, me limitaré a exponer brevemente a algunos pensadores, cuyas contribuciones —y errores— siguen siendo relevantes para nuestros propósitos.

Después de hacer una exposición altamente selectiva de la doctrina leibniziana de la modalidad, tendré algo que agregar acerca de las opiniones de John Stuart Mill y Max Weber sobre el razonamiento contrafáctico en las ciencias sociales e históricas.

Entre Aristóteles y los avances recientes reseñados en el capítulo 1, ciertamente Leibniz es la figura más importante en la historia de la lógica modal. De su trabajo surge la definición de la necesidad como *la verdad en todos los mundos posibles*, así como la analogía entre la lógica deóntica y la lógica modal alética (es decir, la analogía entre la obligación y la necesidad) y la distinción entre la posibilidad y la composibilidad. Sin embargo, esta última distinción está elaborada de una forma que hace que le resulte muy difícil a Leibniz admitir que las afirmaciones contrafácticas tienen significado. Para ver esto, debemos esbozar —nuevamente, en forma muy resumida— parte de las doctrinas metafísicas de Leibniz. Como es bien conocido —pero no comprendido— Leibniz sostenía que el mundo estaba compuesto de *mónadas* o sustancias individuales. Una sustancia individual es completamente descrita a través del correspondiente concepto individual, que incluye todo lo que le ocurrirá a ese individuo y todas las propiedades que le pueden ser adscritas con verdad. Así, todas las afirmaciones verdaderas son analíticas, es decir, el predicado sólo repite lo que ya está incluido en el sujeto. De esta tesis se deduce inmediatamente que las afirmaciones contrafácticas que establecen, sea en el antecedente o en el consecuente, que un individuo del mundo real hubiera actuado de forma diferente o hubiera sido diferente en otro mundo no tienen significado, ya que entonces él no hubiera sido *ese* individuo.

En realidad, la restricción a los contrafácticos impuesta por la teoría de Leibniz es aún más fuerte, como se explicará a continuación. Dada una sustancia posible —es decir una noción completa de un individuo— sabemos que puede ser actualizada, pero no sabemos todavía bajo qué condiciones. No todas las sustancias posibles pueden ser realizadas simultáneamente, es decir, en el mismo mundo; se restringen mutuamente de diversas formas. Un mundo posible es un conjunto máximo de composibles, es decir un conjunto consistente de sustancias posibles tales que la suma de una sustancia más haría que el conjunto fuera inconsistente. Ahora bien, en algunos textos Leibniz habla como si una sustancia pudiera ser miembro de varios mundos posibles,[4] pero Nicholas Rescher ha argumentado de manera convincente que esta idea no es en realidad compatible con la metafísica leibniziana.[5] Debido a que Leibniz argumenta que la noción de una sustancia individual contiene una referencia a todas las demás sustancias en el mundo en el que se realiza, es imposible el trasplante de una sustancia de un mundo a otro. Formalmente, la relación de composibilidad es una relación de equivalencia (reflexiva, simétrica y transitiva) que realiza una partición del conjunto de sustancias abstractamente posibles en clases equivalentes o mundos posibles. Así, no sólo es ilegítimo atribuirle a un individuo acciones o propiedades que no se encuentran correlacionadas

con él en el mundo real, sino que uno ni siquiera puede suponer que él vive en un mundo posible diferente del mundo real en algunos insignificantes detalles, aparentemente no relacionados con ese individuo, tales como la ubicación de una mota de polvo en el otro lado de la luna. O más propiamente, el punto es que no existen detalles insignificantes: no hay forma de modificar un conjunto de sustancias individuales sin modificar cada una de las sustancias.

Esta teoría puede parecer traída de los pelos, y supongo que lo es. Sin embargo, puede servir para destacar las dificultades inherentes a lo que podemos llamar, con un barbarismo,"contrafactibilización ontológica". En la sección siguiente, gran parte de esta discusión rondará alrededor de una distinción entre esta forma de interpretar los contrafácticos y la "contrafactibilización epistemológica". La primera, que también puede llamarse una "teoría de mundos paralelos", supone que cualquier pequeño cambio contrafáctico en algún punto en el tiempo implica la contrafactibilización de todo el mundo previo (y tal vez posterior) a ese momento. En esta versión, los mundos posibles (excluyendo mundos con leyes naturales no deterministas) no tienen puntos en común donde converger o divergir, sino que siempre corren paralelos entre sí. Esta segunda interpretación podría ser considerada una "teoría de mundos ramificados", ya que permite que dos mundos sean idénticos hasta un punto y se diferencien desde ese punto en adelante (y eventualmente converjan, otra vez, más tarde). En la segunda teoría, no hay problema de "identidad transmundo", es decir, de determinar qué individuos son comunes a varios mundos. Aceptando la noción de Kripke y Dummett de que las propiedades esenciales de un individuo son las propiedades que están de alguna manera ligadas a sus orígenes, una condición suficiente para que dos individuos en dos mundos sean el mismo individuo es que fueran el mismo individuo antes del punto de ramificación. En la primera teoría, sin embargo, es difícil ver cómo podría adelantarse un criterio tan preciso de identidad transmundo. Por cierto, uno podría establecer algunos criterios tales como el de la secuencia ADN, pero, al suponer cambios acordes con el medio, dos individuos con la misma secuencia pueden tornarse casi tan diferentes fenotípicamente como uno quiera. Por lo tanto, Leibniz está simplemente cortando el nudo gordiano: del hecho de que en una teoría de mundos paralelos todos los criterios de identidad deben ser arbitrarios, él saca la conclusión de que los individuos están completamente ligados a un mundo y que es imposible trasplantarlos.

Una salida sería sostener que la identidad puede ser una cuestión de grado, como en el caso de la teoría de contrapartes de David Lewis, que también aparece en Leibniz.[6] De acuerdo con esta teoría, un individuo que vive en un mundo no puede también aparecer en otro, pero podemos encontrar una o varias *contrapartes*, es decir personas que se le parecen más o menos en algún aspecto o aspectos. Las proposiciones contrafácticas, entonces, deben ser leídas como si fueran acerca de la(s) contraparte(s) de

la persona en algún mundo posible. Más concretamente, la proposición: "Si Sixto no hubiera violado a Lucrecia, habría tenido una vida feliz", debe ser leída como si dijera que el individuo que (i) no violó a Lucrecia y (ii) que en los demás aspectos es mínimamente distinto del Sixto real también (iii) vivió una vida feliz. Obviamente, esto gira alrededor de la frase"es mínimamente distinto", acerca de la cual habrá más, en la sección siguiente.

No conozco antes de John Stuart Mill ningún tratamiento de contrafácticos en las ciencias sociales. En la primera discusión metodológica de lo que entonces se conocía como las "ciencias morales", Mill rechaza el Método de la Diferencia en las ciencias sociales, sobre la base de que no podemos imaginar una sociedad que cambie en algún aspecto particular mientras que los demás aspectos continúan iguales. Vale la pena citar el párrafo completo:

> Si pudieran encontrarse dos naciones que fueran similares en todas las ventajas y desventajas naturales; cuyas poblaciones se parecieran en cada una de sus cualidades, físicas y morales, espontáneas y adquiridas; cuyos hábitos, usos, opiniones, leyes e instituciones fueran iguales en todos los aspectos, excepto que una de ellas tuviera un impuesto a la importación más alto, o que en otros aspectos interfiriera más con la libertad de industria; si una de estas naciones fuera rica y la otra pobre, o una más rica que la otra, éste sería un experimentum crucis: una verdadera prueba a través de la experiencia sobre cuál de los dos sistemas es más favorable a las riquezas nacionales. Pero la suposición de que se puedan encontrar dos instancias tales es manifiestamente absurda. *Tampoco es posible esta concurrencia ni siquiera abstractamente. Dos naciones que concordaran en todo excepto en su política comercial también concordarían en eso.*[7]

Este argumento es bastante extraño, ya que parecería ser igualmente válido para otros ejemplos del Método de la Diferencia. En su discusión general del método, Mill esconde esta dificultad, concediendo que "no es necesario extender esta similitud de circunstancias a aquellas que ya se sabe no son esenciales para el resultado",[8] una cláusula que es bastante dudosa en sí misma —ya que ¿cómo podemos saber de antemano qué es esencial y qué no lo es?— a la vez que inconsistente con la discusión de las ciencias sociales, en las que se requiere una similitud *total*. Sin embargo, Mill trata de justificar esta inconsistencia, argumentando que la inaplicabilidad del Método de la Diferencia para las ciencias sociales se debe a que la característica distintiva (la interferencia en la industria y el comercio) es aquí "un efecto de causas preexistentes" y no una "propiedad de Tipos".[9] Si, entonces, vamos a la discusión de las propiedades de Tipos,[10] estas son propiedades que pueden ser tomadas como un *índice* de otras propiedades con las cuales están constantemente asociadas, sin ser la *causa* de estas propiedades. Así, la humanidad es un Tipo, porque las propiedades que son específicas del hombre son independientes entre sí, aun si siempre se las encuentra juntas. En contraste, la clase de los cristianos no es un Tipo, porque todas las propiedades específicas de los miembros de esta clase

surgen de su propiedad de ser cristianos. Esto implica que el Método de la Diferencia —es decir, el método comparativo— es inaplicable a los Tipos, ya que los Tipos siempre difieren en más de una característica. Por otra parte, el método contrafáctico es aplicable a los Tipos, porque se supone que las características son independientes entre sí y por lo tanto pueden ser libremente variadas en la imaginación.

En las ciencias sociales, sin embargo, ni el método comparativo ni los experimentos del pensamiento contrafáctico tienen lugar. La razón por la cual Mill considera que los contrafácticos son ilegítimos en el estudio de las sociedades es que aquí todas las partes se suponen independientes a tal punto que ninguna parte o subconjunto de las partes en sí misma puede ser cambiada en la imaginación si la otra se mantiene constante. Acepto que esto es una instancia de una actitud preteórica, porque obvia relacionar la interdependencia con una teoría específica. En el nivel ontológico podemos sostener, sin duda, que todas las cosas de alguna manera tienen consistencia en una vasta red causal, pero la ciencia es una búsqueda de una interdependencia *específica*. Cuando se discute el impacto del libre mercado sobre la prosperidad de una nación, uno tiene que utilizar un marco teórico con un número finito de variables especificadas, y no es para nada absurdo suponer que dos sociedades puedan tener los mismos valores para todas estas variables menos una, ya que esto no implica que sean también similares en todos los demás aspectos no incluidos en la teoría.

Hay una actitud preteórica complementaria que, en lugar de suponer que ningún contrafáctico es legítimo en el estudio de la sociedad, tiende a sostener que todos los contrafácticos son completa e igualmente legítimos. Esta actitud es, creo, la que se encuentra en un famoso artículo de Max Weber sobre la posibilidad objetiva y la adecuación causal en las ciencias históricas. La noción de la posibilidad objetiva es claramente una subclase de la noción de contrafáctico, pero Weber no pone suficiente acento en los aspectos en los que ésta difiere de la posibilidad meramente subjetiva. Antes de llegar a mis comentarios críticos, debo decir que hay mucho que merece ser leído con provecho en el análisis de los contrafácticos de Weber. Una idea importante, a la que retornaré más tarde, es su observación de que la elección de contrafácticos *ex post* es a menudo guiada por la variedad de opciones subjetivas abierta a los actores *ex ante*. Que un acto dado sea caracterizado como una posibilidad objetiva, depende a menudo de si pertenece al repertorio subjetivo de los agentes. La contradicción evidente en la oración anterior, donde los factores subjetivos son incluidos entre los elementos objetivos, desaparece ante una reflexión más cuidadosa. El contraste entre "objetivamente presente en la situación histórica" y "subjetivamente creado por el historiador" no coincide con la oposición entre los "aspectos materiales y estructurales de la situación" y los "elementos psicológicos y culturales de la situación". Obsérvese también que en una teoría de los mundos ramificados la gama subjetiva de posibilidades es una coyuntura natural en la que insertar la cuña de la contrafactibilización, aunque no diría que no hay otros puntos de ramificación además de aquéllos ubicados en los momentos de elección.

Otra percepción importante que brinda el ensayo de Weber es su observación de que el razonamiento y la práctica jurídica es un paradigma para el análisis de la importancia causal de varios factores en una situación histórica compleja. No queda claro, sin embargo, que también tiene razón cuando ve como diferencia entre dos campos que el problema jurídico de la intención subjetiva es de menor importancia para la historia, a la que le interesan los hechos más que quien los hace. Por cierto, el contraste puede invertirse. En el estudio de las contradicciones sociales en el capítulo 5 vimos que hay que darle una importancia decisiva a las intenciones de los actores, ya que sólo a través de ellas podemos comprender cuáles del conjunto de eventos constituyen consecuencias *no* buscadas. Por el contrario, el atribuir culpas subjetivas no siempre es consecuencia del examen de los pensamientos reales del culpable; por lo menos a veces depende de la reconstrucción de los pensamientos contrafácticos que un "hombre razonable" hubiera tenido en la misma situación.

Finalmente, Weber tiene razón al subrayar que en la evaluación de las consecuencias de una presunción contrafáctica debemos depender del conocimiento nomológico más que de nuestra empatía subjetiva. El defecto en el análisis de Weber radica en la ausencia, virtualmente total, de lo que llamaré el problema de la legitimidad: ¿son todas las presunciones contrafácticas igualmente justificadas? En una lectura cuidadosa del texto de Weber, encontré sólo un párrafo[11] que muestra *alguna* conciencia del problema y el principal mensaje parece por cierto ser que cualquier "factor" o "elemento" puede ser variado libremente a voluntad. El simple punto que Weber no tomó en cuenta es éste: las leyes nomológicas que utilizamos para concluir del antecedente el consecuente de una afirmación contrafáctica, también deben ser usadas para asegurar que el antecedente mismo sea compatible con los elementos que *no* se supone que varíen. La teoría que forma el marco del análisis puede dejarnos *alguna* libertad (según Mill) para variar la situación, pero (según Weber) no nos dan libertad completa. Tanto Mill como Weber olvidan la importancia de tener una *teoría específica* que pueda circunscribir el conjunto de presunciones contrafácticas legítimas: Mill, al recurrir a una teoría universal (determinismo metafísico), y Weber, al no apelar a ninguna teoría. Estos errores parecen opuestos, pero son las dos caras de la misma moneda en el modo precientífico. Cuando todo es posible, nada es posible.[12]

El problema de los contrafácticos históricos

A través de los comentarios que anteceden, me he comprometido implícitamente con un análisis especial de las afirmaciones contrafácticas. En esta sección, este análisis será presentado en más detalle y contrastado con el análisis alternativo ofrecido por Robert Stalnaker y David Lewis. Debería dejar en claro desde el comienzo que no pretendo ofrecer un análisis general de los contrafácticos y su lógica. La teoría de Stalnaker-

Lewis es sólo uno de los tantos análisis, que yo creo equivocados, por las razones que doy más adelante. Algunas de estas razones son bastante generales, mientras que otras están relacionadas con la falta de pertinencia de su análisis con respecto al problema específico de los contrafácticos históricos. Trato de argumentar que para el historiador la explicación aquí ofrecida es más satisfactoria, pero no pretendo que esta explicación resuelva el problema general de los contrafácticos. Para poner la cuestión en una forma algo diferente, creo que una teoría general válida de las afirmaciones contrafácticas debe ser compatible con el presente análisis de los contrafácticos históricos, pero no tengo indicios sobre cómo debería hacerse la extensión.

Hay dos distinciones que pueden ser útiles para imponer una estructura a los diversos análisis de afirmaciones contrafácticas ofrecidas. En primer lugar, podemos distinguir entre las teorías *ontológicas* y *metalingüísticas* de los contrafácticos. Una teoría ontológica dice que los contrafácticos son *acerca* de los mundos posibles, y son verdaderos o falsos en virtud de las características de estos mundos, de la misma forma que la afirmación "la nieve es blanca" es considerada verdadera o falsa por la presencia o la ausencia de nieve blanca en el mundo real. Una teoría metalingüística, por otra parte, hace que el elemento explicativo principal sean las *afirmaciones* sobre el mundo más que las características (no lingüísticas) del mundo. A grandes trazos, el análisis metalingüístico del contrafáctico "Si p hubiera ocurrido, entonces q habría ocurrido" se centra alrededor de la existencia de algún conjunto S de afirmaciones empíricas y legaliformes tales que la conjunción de S y p lógicamente implica q. Siguiendo a David Lewis podemos hacer una distinción más dentro de los enfoques metalingüísticos, según que los contrafácticos sean analizados en términos de *condiciones de verdad* o en términos de *condiciones de afirmabilidad*. Una cita parafraseada de Lewis explicará la diferencia, mejor de lo que yo puedo esperar hacerlo:

> La mayor parte de las anteriores teorías de los contrafácticos son *metalingüísticas*: un contrafáctico es verdadero o afirmable si y sólo si su antecedente, junto con las demás premisas apropiadas, implica su consecuente. O el contrafáctico es una oración, lo que quiere decir que existe algún argumento de ese tipo, o es en sí mismo una presentación elíptica de tal argumento. En la versión anterior, el contrafáctico puede ser evaluado como verdadero o falso según que existan o no premisas adecuadas S las que, junto con p, impliquen q. En la versión posterior, no puede ser evaluado como verdadero o falso, sino —después de que las premisas omitidas S hayan sido restauradas— solamente como válido o no válido.[13]

Postulo que al historiador sólo le importan las condiciones de afirmabilidad; por lo tanto, la teoría presentada a continuación es una teoría metalingüística de las condiciones de afirmabilidad de las declaraciones contrafácticas. El punto fundamental es que las afirmaciones contrafácticas

deben ser explicadas en términos de alguna teoría *real T*, que el proponente, si es desafiado, podría producir para respaldar su afirmabilidad. Creo que esto tiene sentido en el tratamiento de los contrafácticos históricos. Es totalmente indiferente en la investigación histórica real la mera existencia de alguna teoría —posiblemente desconocida y tal vez incognoscible— que hiciera válida la inferencia. Pero un problema completamente diferente es si mi análisis en términos de las condiciones de afirmabilidad debiera ser complementado, para el propósito del análisis lógico, por una explicación de las condiciones de verdad. Es decir, ¿puede un análisis, en términos de condiciones de afirmabilidad, ser una explicación *completa* de la lógica de los contrafácticos o sólo funciona como una explicación de cómo algunos de tales contrafácticos son utilizados en realidad en el trabajo científico? Lewis argumenta que debemos disponer de alguna teoría de las condiciones de verdad, porque de otra manera no podríamos encontrarle el sentido a contrafácticos no expresados, que sirven, por ejemplo, como el antecedente de un contrafáctico ordinario. ("Si fuera verdad que la Segunda Guerra Mundial nunca hubiera ocurrido si Hitler no hubiera nacido, entonces habría que fusilar a todos los psicópatas con diagnóstico cierto para impedir futuras guerras.") No estoy seguro de que tengamos intuiciones muy firmes sobre tales casos, y no excluiría *a priori* que, si se hiciera un análisis más cuidadoso, éstas resultaran no tener significado. No es, sin embargo, un análisis de ese tipo el que proponemos aquí. Si pudiera dar un argumento razonado para el punto de vista de que los contrafácticos pueden ser completamente analizados en términos de afirmabilidad, habría dado una alternativa a la teoría de Lewis-Stalnaker. Repito que ésta no es mi ambición. Trato de (i) demoler la teoría de Lewis-Stalnaker de condiciones de verdad en términos de mundos posibles, (ii) ofrecer un análisis de condiciones de afirmabilidad, pero no (iii) probar que es sospechosa la noción misma de condiciones de verdad para los contrafácticos.

Creo que para el análisis de los contrafácticos históricos se necesita una versión más específica de la explicación metalingüística. Introduciendo un aspecto dinámico en la teoría que es parte del conjunto S de los enunciados que permiten la inferencia de p a q, pueden afirmarse algunas cosas que de otra manera no surgirían fácilmente en una explicación más general. Diré entonces que una teoría T explica un proceso P cuando:

(1) El proceso P es en cada momento de tiempo discreto descrito exhaustivamente a través de las variables $x_1 ... x_n$ (x en notación de vector).

(2) La teoría T establece que

(i) para cada variable x_i existe un conjunto X_i de valores individualmente admisibles;

(ii) existe un conjunto X de valores conjuntamente admisibles de las variables, siendo X un subconjunto del producto cartesiano $X_1 X_2 ... X_n$;

(iii) para cada vector admisible $x(t)$ existe por lo menos un vector admisible tal que T excluya que el proceso pueda ser representado por ese vector en $t + 1$.

(3) La secuencia de los vectores $x(t)$ que describen a P es compatible con los requerimientos (2) (i) - (2) (iii).

La condición (2) (ii) es un requerimiento de *compatibilidad estática*. Dice que entre los muchos vectores posibles de manera abstracta que pueden formarse al combinar elementos en el X_i, sólo un subconjunto de vectores puede representar el proceso. Es natural esperar que éste sea un subconjunto adecuado, para que por lo menos una posible combinación de valores no pueda nunca ocurrir como un estado del proceso. La condición (2) (iii) es un requerimiento de *compatibilidad dinámica*. Establece que hay algunas restricciones mínimas sobre los estados que pueden sucederse unos a otros. Creo que estos son los requerimientos absolutamente mínimos de cualquier teoría científica. Están basados en la idea de que el *conocimiento avanza a través de la exclusión*, la reducción del conjunto de configuraciones abstractamente posibles a un conjunto que sea más pequeño, por lo menos en uno.[14] La mayor parte de la gente tiene en mente requerimientos sustancialmente más fuertes cuando se refieren a teorías científicas, a menudo de acuerdo con alguna idea de que el *objetivo del conocimiento es la determinación*. Les gustaría, probablemente reemplazar a (2) (iii) por

(2) (iii)': Para cada $x(t)$ admisible, T excluye a todos los vectores menos uno de X como valores posibles de $x(t+1)$

Este es el enfoque usual para los sistemas dinámicos, en los que uno traza de manera determinista la trayectoria del sistema a lo largo del tiempo. Creo, sin embargo, que para muchos propósitos esto es demasiado restringido. Un argumento general para utilizar (2) (iii) en lugar de (2) (iii)' es el siguiente. Como se dijo con anterioridad en el capítulo 3, cualquier descripción de la realidad debe estar subespecificada, y la subespecificación del presente implica una posible subdeterminación del futuro. En lugar de utilizar modelos deterministas y casi invariablemente inexactos, podemos preferir utilizar uno no determinista y exacto. Con el reciente desarrollo de las herramientas matemáticas para analizar las correspondencias o funciones a las que el conjunto da valores,[15] la última opción debería ser más atractiva de lo que ha parecido tradicionalmente, aunque no sugeriría que debiera emplearse en todos, ni siquiera en una minoría sustancial de casos. Se podrían aducir argumentos más específicos para el uso de funciones con valores conjuntísticos en las ciencias sociales.[16] En el nivel práctico, simplemente faltan teorías que sean tanto deterministas como razonablemente exactas. No estoy diciendo que los hechos sociales están inherentemente menos sujetos al determinismo causal que los hechos naturales, sólo que en este momento debemos tratarlos como si lo fueran, construyendo nuestras teorías para admitir más grados de libertad de lo que es usual en las ciencias naturales. En un nivel más especulativo, pienso que en algunos casos el proceso *debe* ser tratado como si fuera in-

trínsecamente no determinista, aun si sabemos que en la"realidad misma" deba ser completamente determinado.[17]

El vínculo con los contrafácticos puede ahora trazarse de la siguiente manera. Asumimos que en un tiempo *t* la configuración histórica real de los valores de las variables $x_1...x_n$ fue $A_1...A_n$. Entonces, la forma general de un *contrafáctico estático* es el siguiente (suponemos sin falta de generalización que sólo las últimas variables *k* varían independientemente):

Si en un tiempo *t* las variables $x_{n-k+1}x_{n-k+2}...x_n$ hubieran tomado los valores $B_1...B_k$, entonces en el tiempo *t* las variables $x_{i1}x_{i2}...x_{ip}$ habrían tomado los valores $C_1...C_p$, de tal modo que $i_r < n - k + 1$ para todas las *r*.

Podemos hacer dos preguntas diferentes sobre este contrafáctico: con respecto a la significación y a la afirmabilidad. En lugar de significación, hablaré también de la *legitimidad* del antecedente. Como requerimiento mínimo debe haber alguna **x** en *X* de modo tal que $x_r=B_r$ para $r=n-k+1,n-k+2...n$. Esta es la condición de legitimidad estática. Además, podemos querer imponer alguna condición de legitimidad dinámica, por ejemplo que hubiera alguna *k* tal que, del estado real del proceso en *t-k*, hubiera una *trayectoria permitida de estados* (es decir una secuencia de estados que satisficieran la condición 2(iii) antes mencionada), llevando a un estado **x** tal que $x_r=B_r$ para $r=n-k+1,n-k+2...n$. El problema es si demandaremos que un antecedente contrafáctico deba ser capaz de ser insertado en el pasado real. Argumentaré aquí que por cierto éste es un requerimiento natural, pero el lector deberá esperar al fin de esta sección y a los ejemplos de las secciones siguientes para encontrar los detalles del argumento. Con respecto al problema de la afirmabilidad, diferiré la discusión hasta que haya adelantado los elementos esenciales de la teoría ontológica de los contrafácticos.

La forma general de un contrafáctico dinámico es la siguiente:

Si en un tiempo *t* la configuración hubiera sido $A_1...A_{n-k}B_1...B_k$, entonces en un tiempo $t + t'$ el subconjunto $x_{i1}x_{i2}...x_{ip}$ de las variables hubiera asumido los valores $C_1...C_p$.

Nuevamente se podrían hacer las mismas preguntas, recibiendo a grandes rasgos las mismas respuestas. Para que el antecedente sea legítimo debemos imponer el requerimiento estático de que la configuración en *t* sea un miembro de *X*, y (o como explicaré) el requerimiento dinámico de que la configuración sea capaz de inserción en el pasado real. Para el problema de afirmabilidad, una vez más remitiré al lector a las discusiones que se hacen más adelante.

Antes de avanzar hacia una presentación detallada del punto de vista ontológico sobre los contrafácticos, deberíamos notar algunas características de la explicación recién presentada. Un aspecto crucial es que la teoría *T* surge como algo más que simplemente un instrumento que nos permite

concluir del antecedente hipotético el consecuente hipotético: también sirve como un filtro para la aceptación o el rechazo del antecedente mismo. Este doble rol de la teoría T en la evaluación de un contrafáctico crea lo que pienso es *la paradoja básica de los contrafácticos*: mientras más fuerte (es decir, más determinista) sea la teoría *T*, mejor basada estará la conclusión de antecedente a consecuente, pero más vulnerable será también la legitimidad del antecedente. Con una teoría débil, muchos antecedentes son permitidos por el filtro, pero puede ser imposible probar la afirmbilidad del condicional. Así, para que un análisis contrafáctico sea exitoso hay que establecer un delicado equilibrio: la teoría debe ser lo suficientemente débil para admitir la presunción contrafáctica y lo suficientemente fuerte para permitir una conclusión clara. Para utilizar un ejemplo a ser discutido más tarde: cuando evaluamos las consecuencias de una economía sin ferrocarriles en los Estados Unidos del siglo XIX, no podemos preguntar si el motor de combustión interna hubiera sido inventado antes de lo que en realidad lo fue, ya que esto requeriría una teoría del cambio técnico que podría impedir que el mero asumir la idea de la falta de ferrocarriles fuera significativa. El problema no es que es difícil saber si el motor de combustión interna hubiera sido inventado,[18] sino que, si fuera posible contestar la pregunta, no se la habría puesto en primer lugar .

La paradoja básica de los contrafácticos puede también ser llamada como *la falta de importancia de lo inevitable*. Para evaluar la importancia causal de un evento histórico o un hecho social, la remoción mental de ese elemento debe ser legítima, estática y dinámicamente; inversamente, cualquier elemento que no pueda ser sujeto a este experimento del pensamiento es inherentemente imposible de ser sometido a la evaluación causal. Un ejemplo muy diferente, que arroja luz sobre esta dificultad, puede ser tomado del análisis conceptual del poder. Alvin Goldman[19] ha argumentado que, "Decir que *S* es poderoso no es decir que a menudo consigue lo que quiere, sino que cualquier cosa que él quisiera la podría conseguir, sea lo que fuere lo que quisiera." La razón por la que queremos imponer esta condición es que de otra manera deberíamos tener que atribuir poder extremo al *camaleón*, es decir la persona que siempre cambia sus preferencias (sinceras) para estar siempre del lado ganador. Por otra parte, creo que esta contrafactibilización del poder nos lleva a la paradoja básica que acabamos de discutir. Tomemos los dos ejemplos contrastantes de

—Nixon tenía el poder para terminar la guerra de Vietnam: por lo tanto la guerra de Vietnam podría haber continuado si Nixon lo hubiera querido.

—De Gaulle tenía el poder para proteger la independencia de Francia: por lo tanto Francia hubiera perdido su independencia si De Gaulle lo hubiera querido.

La presunción contrafáctica incluida en la afirmación sobre Nixon parece perfectamente legítima; no hay nada inherentemente improbable, por lo que sabemos, en la idea de que Nixon quería que la guerra de

Vietnam continuara. (Esto no quiere decir, por supuesto, que es afirmable el contrafáctico en su conjunto.) Por otra parte, me parece claramente absurdo asumir que De Gaulle *pudiera* haber deseado que Francia perdiera su independencia, aun si queremos atribuirle a De Gaulle el poder que sentimos que tenía. Por cierto, podemos suponer un mundo contrafáctico donde De Gaulle fuera una clase de persona por completo distinta, la clase de persona que hubiera podido desear que Francia perdiera su independencia, pero *ese* De Gaulle probablemente no hubiera tenido ningún poder, porque el poder del De Gaulle (real) surgía principalmente del hecho de que visiblemente no era el tipo de persona que *podría* tener ese deseo. Así, estamos ante un dilema: aceptamos que De Gaulle era por cierto una persona poderosa; aceptamos la necesidad de un análisis contrafáctico de las afirmaciones de poder; pero la afirmación contrafáctica sobre el poder de De Gaulle es o absurda (ya que tiene un antecedente ilegítimo) o falsa (si asumimos un De Gaulle que hubiera deseado otras cosas). De Gaulle, según parece, no tenía poder porque no hubiera podido desear actuar de una forma diferente de la que lo hizo. No veo, por el momento, ninguna salida de este dilema.

El doble rol de la teoría como un filtro y como una maquinaria deductiva señala una importante diferencia entre las afirmaciones contrafácticas y las afirmaciones fácticas. Cuando surgen nuevas teorías además de las viejas (hago abstracción de las nuevas teorías que *reemplazan* a las viejas) el impacto sobre el conjunto de afirmaciones verdaderas (fácticas) causales irá sin ambigüedades en la dirección de un aumento, pero es más difícil de ver el impacto sobre un conjunto de contrafácticos afirmables. Algunos contrafácticos antes aceptados ahora aparecerán como ilegítimos, y algunos que antes parecían legítimos pero imaginables, ahora se podrán afirmar. Sin embargo, esto se mantiene solamente en el corto plazo. Si imponemos alguna condición dinámica de legitimidad, sabemos que el surgimiento de más y mejores teorías, en *última instancia*, reducirá el conjunto de contrafácticos afirmables, ya que en el límite —una teoría totalmente abarcadora y determinista del universo— no existirán puntos de ramificación en los que se pueda insertar la cuña de la contrafactibilización.

Tomaré ahora la teoría de los contrafácticos de Stalnaker-Lewis, que denominamos "la teoría ontológica de los contrafácticos".[20] Como generalmente ocurre en la lógica modal, esta teoría supone que en relación con cualquier mundo dado, incluyendo el real, existe un conjunto de mundos posibles. Además, y éste es el punto fundamental, suponemos que hemos definido alguna medida para la *distancia* entre el mundo dado y los que son posibles en relación con él. Con una ligera simplificación del argumento de Lewis, podemos entonces decir que "si p hubiera sido el caso, entonces q habría sido el caso" es cierto en el mundo x si (1) p no es el caso en ningún mundo que sea posible en relación con x, o (2) q es el caso en el mundo más cercano al nuestro en el que p es el caso.[21] Esto no excluye que pueda haber mundos donde p sea el caso y q no sea el caso; el análisis sólo implica que tales mundos están siempre más lejos del nuestro que, por lo menos, un mundo donde p y q sean ambos el caso.

Quiero elogiar la idea general de esta explicación y criticar la interpre-

tación particular (ontológica) de la noción de mundo posible utilizada por Lewis. Con respecto al elogio, es por cierto importante señalar que para que "si *p* hubiera sido el caso, entonces *q* habría sido el caso" sea cierta (o afirmable), no deberíamos requerir que q hubiera sido el caso en *todas* las circunstancias en las que *p* hubiera sido el caso. Argumentar que si Hitler hubiera muerto justo en el momento de nacer, la Segunda Guerra Mundial no hubiera ocurrido, no nos compromete con la proposición, "Si Hitler hubiera muerto justo en el momento de nacer y si hubiera también tenido un mellizo idéntico que sobreviviera y creciera, la Segunda Guerra Mundial no habría ocurrido." Lo que queremos decir es que la Segunda Guerra Mundial no habría ocurrido si suponemos la temprana muerte de Hitler y *un mínimo de otros cambios*, entre los que asumir la existencia de un hermano mellizo probablemente esté más allá de tal mínimo. Con respecto a lo que *es* exactamente el mínimo, diría que en la teoría metalingüística de las condiciones de afirmabilidad no hay necesidad de suponer ningún cambio además de la remoción mental de Hitler, por lo menos para el propósito de explorar las consecuencias de la Segunda Guerra Mundial. Esto es así porque en lo relativo a una teoría socioeconómica en el macronivel, la existencia de un individuo en especial es un accidente o un hecho exógeno. El impacto de un individuo de un tipo dado sobre un proceso social es tema de una explicación sociológica, pero no el surgimiento de ese individuo, a menos que pueda demostrarse que como tipo era característico de su época y no sólo una desviación idiosincrática.

No son tan simples las cuestiones en la teoría ontológica de las condiciones de verdad para los contrafácticos. Asumir la muerte de Hitler en el momento de nacer requiere —*aunque no de una manera única y definida*— una cantidad de otros cambios concomitantes. Puede asumirse que la muerte de Hitler tuvo lugar de una cantidad de maneras diferentes, cada una de las cuales requeriría alguna forma de reelaboración del pasado, y puede ser muy difícil decir cuáles de estas formas implica el mínimo de cambio. En síntesis, sostengo que la relación "implica un cambio mayor que" es sólo un ordenamiento parcial, mientras que Lewis debe sostener que es completo. O sea, si para dos "mundos posibles", *x'* y *x"*, es el caso que *x"* es diferente del mundo real *x* en todos los aspectos en los que *x'* es diferente de *x*, pero no viceversa, entonces podemos decir sin ambigüedades que *x"* implica un cambio mayor que *x'*. Si, por otra parte, hay algunos aspectos compartidos por *x* y por *x'* pero no por *x"* y algunos aspectos (diferentes) compartidos por *x* y por *x"* pero no por *x'*, entonces no es posible ningún ordenamiento no ambiguo. La respuesta de Lewis a esta objeción obvia es invocar una noción de *importancia* tosca pero (en su opinión) no completamente arbitraria que puede ser utilizada para sopesar los diferentes aspectos. Es un poco difícil refutar esta respuesta, ya que no estoy seguro de que la haya comprendido. Sin embargo, pienso que la siguiente consideración tiene algún peso: si el objetivo del contrafáctico histórico es llegar a una evaluación de la importancia relativa de las causas, entonces no podemos utilizar nuestros sentimientos intuitivos sobre la importancia

como una base para sopesar las variables en una comparación entre mundos posibles. No podemos explicar la causación por medio de contrafácticos y luego analizar éstos a través de una noción de importancia, que inevitablemente incluye la importancia causal.

Además de este problema de circularidad me gustaría llamar la atención sobre una dificultad en la teoría de Lewis, que está relacionada con la naturaleza de cuerpo completo (aunque sea brumosa) de sus mundos posibles. De acuerdo con la interpretación delineada aquí (y antes en el capítulo 3), un mundo posible es un simple esqueleto de un estado, descrito a través de un número finito de variables. Para utilizar un concepto explorado en el capítulo 4, *es posible por todo lo que creo (teóricamente)*, una configuración de variables compatible con una teoría dada *T*, que creo verdadera. Sin embargo, los mundos posibles de Lewis son entidades corpóreas sin espacios en blanco a ser llenados. Si tienen un espacio tridimensional (algunos no lo tienen), entonces cada punto en el espacio está o vacío o lleno con alguna materia específica. Ahora, esta existencia sólida (o subsistencia, o lo que sea) de los mundos posibles hace difícil ver cómo uno podría construir una medida de distancia dándole pesos diferentes a características diferentes. Este procedimiento tiene significado para un esqueleto de mundo con un número finito de características, pero no para los mundos compactos de la teoría de Lewis, por lo menos, si aceptamos la afirmación de Ernest Nagel de que "no es posible que ninguna afirmación finita en su longitud pueda formular la totalidad de las características corporizadas en cualquier cosa que tenga existencia concreta".[22] Hay demasiados aspectos para que podamos contarlos, y mucho menos atribuirles peso.

Un problema similar surge en el análisis de Fabricio Mondadori (directamente inspirado por Lewis) sobre el rol de los contrafácticos en la filosofía de Leibniz. Como antes dije, sugiere que una proposición contrafáctica sobre Sixto incluye la noción de un individuo que satisface una cierta condición y que "difiere en lo mínimo" del Sixto real, noción que a grandes trazos se explica como "tiene más propiedades en común con el Sixto real que cualquiera de sus contrapartes".[23] Pero es bien conocido, de acuerdo con Leibniz, que un concepto individual completo sólo puede ser resuelto "al infinito" en una cantidad infinita de propiedades simples. Si varias de las "contrapartes" de Sixto que satisfacen esta condición compartieran una cantidad infinita de sus propiedades y también difirieran en una cantidad infinita de propiedades, mientras que no haya contraparte que satisfaga la condición y que difiera de Sixto solamente en una cantidad finita de propiedades, entonces la explicación puede no estar bien definida.

Es instructivo mirar algunos de los problemas que Lewis dice haber resuelto, y ver cómo estas soluciones se aplican a problemas históricos. Entre las dificultades superadas tanto por la teoría de la regularidad de la causación como por muchas tradicionales teorías contrafácticas, los problemas de los epifenómenos y de las causas potenciales preexistentes[24] parecen especialmente importantes. El primero se refiere a la siguiente

situación: un hecho c es causa suficiente (dadas algunas de las circunstancias reales) para el suceso *e* y causa necesaria (dadas circunstancias distintas de las reales) para el suceso *f*. (También suponemos que *e* precede a *f* en el tiempo.) Entonces, podemos inferir que si *e* no hubiera ocurrido, *f* no habría ocurrido y entonces parecería que *e* es la causa de *f*. Según Lewis, esta confusión entre la correlación y la causación se puede disipar si se supone que, "Deshacerse de *e* conservando invariable a *c* y descartando otras leyes y circunstancias en virtud de las cuales *c* no podía dejar de causar *e*, significa desviarse menos de la realidad que si uno se aferra a esas leyes y circunstancias y descarta a *e* mediante un paso atrás para abolir su causa *c*".[25] Efectivamente, tiene que ser así si la teoría de Lewis es correcta, pero uno no puede estar a favor de ésta o de cualquier otra teoría señalando los problemas que nos permitiría resolver si fuera correcta.[26] Si Lewis pudiera darnos criterios *independientes* de lo que constituye un "gran" alejamiento de la realidad, entonces podríamos constatar si la afirmación citada se mantiene en cualquier caso especial, pero ningún criterio de ese tipo ha aparecido todavía.

T.A. Climo y P.G.A. Howells[27] han aplicado la teoría de Lewis a la causación histórica. El problema del epifenómeno es ejemplificado por un caso en el que *c* es una disminución de la demanda y de los precios de mercaderías británicas en el extranjero y *f* un aumento en el desempleo. El problema, igual que antes, es cómo podemos evitar aceptar la afirmación "si *e* no hubiera ocurrido, entonces *c* no habría ocurrido" la que, dada la verdad de "si *c* no hubiera ocurrido, entonces *f* no habría ocurrido" implica que "si *e* no hubiera ocurrido, entonces *f* no habría ocurrido" y por lo tanto "*e* causó *f*." Los autores tratan de resolver este problema en el siguiente pasaje, que cito en extenso porque lo encuentro difícil de entender:

> La caída en la demanda de exportaciones *c* es sólo uno de varios hechos alternativos que, se puede decir que causaron *e*, la caída en precios (*c* es condición suficiente para *e*). Sin embargo...dado que la demanda interna continúa subiendo, la única otra causa posible de *f*, el aumento del desempleo, está ausente. Por lo tanto, el aumento del desempleo podría haber sido causado sólo por la caída en la demanda de exportaciones. Por lo tanto, si se nos preguntara *si no e entonces no c,* o *si no e entonces c* se aparta menos de la realidad, nos sentiríamos con derecho a decir que la segunda es la que lo hace. El criterio utilizado aquí es la fuerza del vínculo entre la caída en la demanda y la caída en los precios comparados con la fuerza del vínculo entre la caída en las exportaciones y el desempleo. Por cierto, deberíamos decir que es apartarse menos de la realidad abandonar la caída en la demanda de las exportaciones como la causa de la caída de los precios y buscar alguna causa alternativa —tal vez, la política deflacionaria del gobierno— que sostener en este caso los precios bajos como causa del aumento del desempleo.[28]

Mientras que la solución de Lewis tenía por lo menos la virtud de una simplicidad gordiana, esta explicación parece hacer que el nudo sea más apretado de lo que era antes. Aunque la intención de los autores sea aplicar

la teoría de Lewis a los problemas históricos, su análisis difiere del suyo en, por lo menos, tres aspectos.

(1) En la teoría de Lewis no es una cuestión de *si no e entonces no c* o *si no e entonces c* difiere menos con la realidad, sino de si *c, pero no algunas otras leyes y circunstancias*, o *no c* se aparta menos del mundo real.

(2) En la explicación de Lewis el problema del epifenómeno se resuelve sin recurrir para nada al efecto *f*, mientras que Climo y Howells centran fundamentalmente su explicación en hacer que la fuerza del vínculo entre *c* y *f* sea relativamente más débil que el vínculo entre *c* y *e*, porque el segundo representa una condición causal que es suficiente y (dado el supuesto de una subida en la demanda interna) necesaria, mientras que el primero sólo es una causa suficiente.

(3) De acuerdo con Lewis, la alternativa que implica del menor alejamiento de la realidad no implica la sustitución de *c* por una nueva causa *c'*, ni tampoco entra en su análisis de ninguna forma un sustituto de ese tipo. En el mundo real, *c* causó *e*; en el mundo más cercano posible no-*e*-nada causó *e*; es difícil ver por qué deberíamos invocar un mundo en el que *c'* causó *e* y *c* causó *f*. A falta de más clarificación (puedo haber interpretado mal los argumentos) debemos concluir que en este punto algo salió mal en el análisis de Climo- Howells.

El problema de las causas potenciales preexistentes es el siguiente: "Supongamos que c_1 ocurre y causa *e*; y que c_2 también ocurre y no causa *e*, pero hubiera causado *e* si c_1 hubiera estado ausente...¿En virtud de qué diferencia es que c_1 y no c_2 causa *e*?"[29] El problema es que aquí una afirmación verdadera sobre la causalidad (c_1 causó *e*) parece implicar un contrafáctico falso (si c_1 no hubiera ocurrido, entonces *e* no habría ocurrido). A grandes trazos ésta es la solución de Lewis. Supongamos que c_1 causó *e*, a través de la causa intermedia *d*. Entonces, es posible sostener que c_1 podría no haber causado *d*, pero todavía impedido que c_2 causara *e*, de tal modo que *e* no habría ocurrido. Hay en este análisis una ambigüedad. Cuando Lewis escribe que "gracias a c_1 no hay cadena causal de c_2 a *e*", esto puede significar que c_1 le impide a c_2 ser la causa de *e* por alguna interferencia real en la eficiencia causal de c_2 o que c_1 se lo impide a c_2 por el simple hecho de que *e* no puede ser causada dos veces. Tenemos este caso si c_1 es una bala que llega al corazón de un hombre en un tiempo t_1 y c_2 es otra bala que llega al corazón del mismo hombre en un tiempo posterior t_2. Si, para dar un ejemplo fantasioso, la bala c_1, antes de llegar al corazón del hombre, también pasara por el corazón de otro hombre, que al caerse bloqueara el camino entre la bala c_2 y el primer hombre, tendríamos el primer caso. A lo sumo, la explicación de Lewis es adecuada para este ejemplo fantasioso. Además de su falta de adecuación a ejemplos de otro tipo, la teoría está plagada de las dificultades generales antes discutidas.

El ejemplo de Climo-Howells es el siguiente: c_1 es la Segunda Guerra Mundial, c_2 el aumento de empleo a niveles previos a la guerra, *e* la suba de

precios durante la guerra, *d* la reducción en suministros importados y p_a el paquete de controles del gobierno dirigidos a limitar la demanda y racionar los suministros. Si c_1 no hubiera ocurrido, c_2 —a través del intermediario de un aumento del poder adquisitivo— hubiera causado *e*, pero c_1 también causó p_a, que en el mundo real impidió que la intermediación fuera posible. (Así, tenemos un caso análogo al fantasioso ejemplo del párrafo anterior.) Los autores proceden rechazando el contrafáctico *si no d entonces no* c_1 *y no* p_a sobre la base de que "es más fácil concebir un mundo cercano al nuestro en el que hubo una guerra pero en el que por una u otra razón las importaciones no fueron reducidas, que uno en el que las importaciones fueron mantenidas, pero en el que no hubo guerra".[30] Esta afirmación parece centrarse sobre la importancia masiva (¿de acuerdo con cuál criterio?) de una guerra, comparada con una simple reducción de importaciones, pero esto es seguramente una característica accidental de este ejemplo. Sería fácil construir un ejemplo formalmente similar donde *d* fuera la guerra y c_1 alguna característica intrínsecamente insignificante" (lo que quiera que eso signifique), tal como el largo de la nariz de Cleopatra. Nuevamente debo concluir que la explicación de los mundos posibles no tiene poder de resolución cuando se la aplica a casos concretos.

Estoy ahora en mejor posición para completar mi propia explicación de los contrafácticos históricos. Esta explicación depende principalmente de la noción de legitimidad dinámica. Seguiré a Lewis en la estructura formal de la teoría, apelando a los mundos posibles y la distancia entre mundos, pero trataré de hacerlo de una manera no abierta a las objeciones especificadas anteriormente. Ahora, con una teoría *T*, una noción aparentemente plausible de mundos relativamente posibles podría ser la siguiente: un estado s'_t es posible en relación a s_t si en la historia pasada de s_t hubiera un estado s_{t1} tal que hubiera una trayectoria permitida de s_{t1} a s'_t; es decir, si hubiera algún punto de ramificación a partir del cual el proceso pudiera divergir hacia s_t o s'_t. Esta definición funciona bastante bien si s_t es un estado real de la historia del mundo, ya que conocemos el pasado de tales estados. Pero no funciona si s_t mismo es un mundo no realizado, ya que los estados en general no son dados con los pasados que los llevaron hasta ese punto. Esta dificultad podría resolverse,[31] pero en lugar de eso rogaré que sea ignorada. El historiador, seguramente, está interesado en la afirmabilidad de los contrafácticos emitidos en el mundo real; es presumible que no le llame la atención la afirmabilidad de los contrafácticos emitidos en algún mundo posible. Para determinar la importancia relativa de causas en el mundo real, debe postular un mundo en el que una o más de ellas esté ausente, pero no está obligado a averiguar también cuán importantes son las diversas causas en *ese* mundo.

Ahora bien, el criterio dinámico de legitimidad permite una medida natural de la distancia entre el mundo real y algún estado posible: si en la historia pasada del sistema real hubiera algún estado s_{t1}, del cual trayectorias permitidas llevaran tanto al estado real s_t como al estado hipotético s'_t, pero ninguna trayectoria permitida desde s_{t1+1} hasta el estado s'_t, entonces

$(t-t_1)$ es la distancia entre el mundo real y el mundo alternativo en cuestión. Es decir, mientras más atrás tengamos que ir para insertar el estado posible en la historia real, mayor es la distancia con ese estado. Un estado de la economía con el doble del producto nacional bruto de la economía real es posible, pero sólo a la distancia, porque tendríamos que ajustar los ahorros y las decisiones de inversión durante un período muy largo en el pasado para que ese estado estuviera al final de una trayectoria permitida. Utilizando esta medida de distancia también podemos definir de manera natural las condiciones de afirmabilidad de los contrafácticos históricos: el contrafáctico en su conjunto es afirmable si el consecuente vale en el mundo (s) más parecido, donde vale el antecedente. Si, por ejemplo, el antecedente pudiera ser insertado en el mundo real en t_2, mientras debemos ir para atrás a un mundo anterior t_1 para encontrar un punto de ramificación, desde el cual una trayectoria permitida lleve a un estado donde se den tanto el antecedente como el consecuente, entonces el contrafáctico no es afirmable. Tomemos la afirmación:"Si no hubiera sido por la esclavitud, el producto bruto del Sur de Estados Unidos en 1860 hubiera sido el doble de lo que era realmente." Esta afirmación no sería sostenible si un Sur no esclavista pudiera surgir de una ramificación, digamos no anterior a 1750, mientras que un producto bruto del tamaño requerido demandaría cambios contrafácticos que llegaran hasta el año 1700.

Deberíamos agregar una breve definición de las condiciones de afirmabilidad de los *contrafácticos-podría [might counterfactuals]*. La afirmación "si p hubiera sido el caso, entonces q podría haber sido el caso" es afirmable si hay *por lo menos un mundo* x tal que (i) p se da en x, (ii) q se da en x, y (iii) no hay un mundo más parecido que x donde se de p. El requerimiento correspondiente para los contrafácticos-sería [*would counterfactuals*] es que la condición (ii) se de en *todos* los mundos x que satisfacen (i) y (iii). Es decir, la distinción entre contrafácticos-podría y contrafácticos-sería exige la posibilidad de varios mundos igual y mínimamente distantes.

El problema crucial en esta explicación es la necesidad de un criterio dinámico de legitimidad, la exigencia de que el estado alternativo sea pasible de inserción en el pasado real. Mi argumento principal derivará de los ejemplos concretos analizados en las siguientes secciones, y aquí sólo puedo dejar establecido que la preocupación del historiador por el pasado real hace que le resulte difícil aceptar alternativas que no podrían (por lo menos en tanto lo que él sabe) haber ocurrido como resultado final de algún proceso comenzado en algún punto del pasado real. Para imaginarse una economía norteamericana sin esclavitud en 1860 no es suficiente convertir a todos los dueños de esclavos en sus equivalentes capitalistas o semifeudales y a todos los esclavos en jornaleros o aparceros o cualquier forma que se piense que sea la más parecida, y dejar el resto sin cambiar. Aun si esta alternativa fuera estadísticamente legítima, sabemos que no podría haber existido, salvo sobre premisas que difirieran enormemente del período real previo a 1860, pero entonces no estaríamos ya tratando la

esclavitud de los negros norteamericanos.

La idea principal detrás de este argumento puede tal vez presentarse mejor tomando el ejemplo más simple de un individuo con sus propiedades reales y contrafácticas. Algunas de estas propiedades estarán estrechamente vinculadas a su constitución genética, aun si fueran intrínsecamente sin importancia: las huellas digitales pueden ser un buen ejemplo. Otras propiedades serán independientes de la secuencia del ADN, pero serán masivamente importantes para la personalidad del individuo: podemos pensar en varias características personales que fueran sólo apenas restringidas por factores genéticos. De acuerdo a la explicación de Lewis supongo que una persona sin mis huellas digitales pero similar en todos los demás aspectos es "más parecida" a mí que una persona con mis huellas digitales pero sin mi arrogancia o mis celos o lo que fuera. De acuerdo con la explicación de Kripke-Dummett, que discutí en los capítulos anteriores, éste no sería el caso, ya que alguien sin mis huellas digitales no podría haber nacido de mis padres, dada su conformación genética. El problema es que Lewis no me busca a *mí* en otros mundos, sólo a una persona que más o menos se me parezca. Según el análisis de Knipke-Dummett, las afirmaciones sobre lo que yo hubiera hecho en varios casos hipotéticos son afirmaciones sobre mí, que requieren identificación plena y no sólo parcial. Acepto que ésta es también la forma en que el historiador tiende a ver a las afirmaciones contrafácticas: son afirmaciones sobre la economía norteamericana, la única, la que conocemos; no afirmaciones sobre alguna economía que difiere mínimamente de la real.

Imperialismo y colonialismo

En lo que resta del capítulo iré de la teoría a las aplicaciones y discutiré en detalle unos cinco ejemplos de razonamiento contrafáctico en historia. Los dos primeros representan una búsqueda del *mejor pasado* o *del contrafáctico óptimo*: ¿podrían el desarrollo de Africa y la industrialización de Gran Bretaña haber tenido lugar con un costo más reducido? Los últimos tres tienen que ver con los contrafácticos como una herramienta en un análisis positivo más que en uno normativo: ¿cuál fue la importancia causal de las Leyes de Navegación para la Revolución Norteamericana, de los ferrocarriles para el crecimiento económico norteamericano, de la esclavitud para el desarrollo del Viejo Sur? Espero que estos estudios de caso serán de interés por derecho propio, además de permitir una mayor clarificación y sustanciación de las tesis adelantadas en la sección anterior.

Mi primer caso concierne a las consecuencias —y las alternativas— de la colonización europea de Africa. Este no es un campo con el cual esté muy familiarizado, por lo que dependeré en alto grado de un estimulante trabajo de Jarle Simensen.[32] El problema básico discutido en la literatura citada por Simensen es el siguiente: ¿podría la historia de Africa reciente haber tomado otro curso que el presente, y cuáles hubieran sido las

consecuencias de la modernización y el desarrollo económico del continente? Sobre la base de la tipología de las principales alternativas hipotéticas de Patrick Manning presentadas en el debate,[33] Simensen enumera las siguientes posibilidades:

(1) *Desarrollo autónomo* = modernización + autonomía económica + independencia política (ejemplo: Japón).
(2) *Desarrollo periférico independiente* = modernización + desarrollo económico periférico (es decir, dependiente del mercado mundial) + independencia política (ejemplo: Etiopía).
(3) *Colonialismo benigno* = modernización + autonomía económica + dominio extranjero.
(4) *Estancamiento periférico* = estancamiento + desarrollo periférico + independencia política (ejemplo: algunas sociedades africanas asoladas por el comercio de esclavos alrededor de 1850).
(5) *Estancamiento autónomo* = estancamiento + autonomía económica + independencia política (ejemplo: algunas sociedades africanas antes de tener ningún contacto con Occidente).

En cambio, se dice que el verdadero curso de los acontecimientos se caracterizó por la modernización, la dependencia económica y el dominio extranjero. Para la discusión de estas alternativas es fundamental la distinción entre imperialismo económico y colonialismo político, y especialmente el hecho de que el primero ocurriera varias décadas o aún más que el segundo. Así, dada la integración de muchas sociedades africanas al mercado mundial, especialmente a través del comercio de esclavos, la dominación política puede haber sido una característica necesaria para la modernización de Africa; lo que los europeos habían destruido sólo los europeos lo podían reconstruir. Si, por otra parte, vamos más atrás en el tiempo, la posibilidad de una modernización autónoma se hace más realista. De acuerdo con Simensen, "Los críticos del dominio colonial...tienden a emplear una perspectiva a más largo plazo que los apologistas, yendo más hacia atrás para encontrar puntos donde cursos de desarrollo hipotéticos completamente diferentes pueden ramificarse".
Veamos ahora los siguientes contrafácticos:

(A) Si el país *X* no hubiera estado bajo dominio colonial en 1920, no se hubiera modernizado.
(B) Si el país *X* no se hubiera integrado al mercado mundial en 1920, no se hubiera modernizado.

Sobre la base del análisis propuesto en la sección anterior, estos contrafácticos deberían explicarse de acuerdo con lo siguiente:

(A') Si el tiempo t_A es la última fecha en la que hay una ramificación

a partir de la cual una trayectoria permitida lleva a un *X* no colonizado en 1920, entonces todos los estados no colonizados de *X* al final de una trayectoria que se ramifica en t_A o más tarde no se han modernizado.

(B') Si el tiempo t_B es la última fecha en la que hay un punto de ramificación a partir del cual una trayectoria permitida lleva a un *X* no integrado en 1920, entonces todos los estados integrados de *X* al final de una trayectoria que se ramifica en t_B o más tarde no se han modernizado.

En la explicación propuesta por Lewis, presumiblemente el análisis procede así:

(A") Entre los estados no colonizados del país *X*, el (los) estado(s) que tienen la mayor similitud de conjunto con el real *X* de 1920 no está(n) modernizado(s).

(B") Entre los estados no integrados del país X el (los) estado(s) que tienen la mayor similitud de conjunto con el X real de 1920 no está(n) modernizado(s).

Si aceptamos la explicación histórica delineada anteriormente, (*A*) es afirmable y (*B*) no lo es simplemente porque (*A'*) es cierto y (*B'*) no lo es. Por otra parte, no parece existir ninguna razón por la que (*A"*) debería ser cierto y (*B"*) falso. Tal vez uno pudiera argüir que se podría hacer que un mundo sin colonización fuera consistente con las leyes conocidas (suponiendo que quisiéramos retener*las*), ya sea renunciando a la modernización y manteniendo todo lo demás intacto o reteniendo la modernización y haciendo los cambios necesarios en los subsistemas educativo y administrativo, pero no puedo ver por qué el primer ajuste debería ser más pequeño que el segundo, como lo hubiera tenido que ser para que (*A"*) explicara a (*A*). También me parece que (*B"*) es más verdadero que falso. Presumiblemente la manera más simple de dejar de lado la integración al mercado mundial es remover todas las actividades de exportación e importación y reubicar a los trabajadores en el comercio o la manufactura local, lo que de seguro también acabaría con la modernización. En un país autónomamente modernizado (comenzando en t_B) los sectores modernos de la economía no necesariamente estarían ligados al comercio extranjero, pero en el mundo real la modernización y la integración al mercado mundial estuvieron tan entrelazadas que abandonar la segunda y retener la primera nos llevaría muy lejos del estado real. Obviamente, no me siento muy confiado al hacer estas afirmaciones, ya que presuponen una noción de similitud de conjunto cuya utilidad y aun significación rechazo; sin embargo, me parece que son el tipo de afirmaciones que un defensor de la teoría de Lewis debería tener que aceptar.

En este contexto surgen otros problemas conceptuales. Es natural, por ejemplo, preguntarse cuál de los estados alternativos antes enumerados son por completo plausibles. El *colonialismo benigno*, que presumiblemente hubiera llevado a un máximo de modernización con un mínimo

de costos, es rechazado por completo por Manning como una "contradicción en los términos", que puede por cierto ser una expresión apta si la frase es entendida en un sentido absoluto (compárese el análisis de la contradicción entre el amo y el esclavo, en el capítulo 4). Pero algunas políticas coloniales por cierto fueron más benignas que otras, o por lo menos surgieron de intereses de largo plazo de los colonizadores más que de sus cálculos miopes de corto plazo. Puede entonces valer la pena discutir si alguna política alternativa podría haberse seguido y, si se hubiera hecho, cuáles hubieran sido las consecuencias. Un criterio posible (sugerido por Simensen) podría apelar a las políticas que fueron consideradas realistas por contemporáneos bien informados y en realidad propuestas por algunos grupos interesados. La discusión de la posibilidad política del capítulo 3 ha demostrado, sin embargo, que este criterio es al mismo tiempo demasiado amplio y demasiado estrecho. Es demasiado amplio porque no permite la posibilidad de que aun los observadores bien informados puedan haber estado equivocados, y demasiado estrecho porque oscurece los casos ambiguos o fronterizos en que no es posible hacer una evaluación *ex ante* de la posibilidad política. Sin embargo, como se observó en la discusión anterior de Max Weber, estas posibilidades subjetivas constituyen una cuña natural para la contrafactibilización.

Si —o en la medida en que— el "colonialismo benigno" es una contradicción en los términos, éste no tiene nada que hacer con nuestro conocimiento sobre la nación colonizada; la imposibilidad surge de nuestras suposiciones acerca del colonizador. Esta distinción se hace crucial cuando evaluamos lo plausible de una modernización autónoma (o de un estancamiento autónomo permanente). Desde el punto de vista de las naciones europeas, la expansión continua y, en última instancia, el contacto con los países africanos era inevitable, lo que quiere decir que es ilegítimo suponer que los segundos podrían haberse mantenido fuera del mercado mundial. Sin embargo, desde el punto de vista de los países africanos no había ninguna necesidad inherente que los llevara a la integración al mercado mundial. En este sentido, el contrafáctico "si Africa no se hubiera encontrado con Occidente, habría experimentado una modernización autónoma" es legítimo, mientras que el contrafáctico "si Occidente no se hubiera encontrado con Africa, la industrialización europea se habría retardado" no lo es. Esta aparente paradoja se resuelve cuando recordamos que el objeto de estudio depende de nuestra elección. Si se define al "sistema" como el continente africano y el resto del mundo es simplemente considerado un medio constante, entonces la expansión de ese sistema en el medio no se sigue de lo que asumimos teóricamente respecto del sistema. Si, por otra parte, Europa es el sistema de variables considerado, entonces la expansión en el medio se sigue inmediatamente de un análisis del capitalismo en términos de reinversión de las ganancias. Desde el punto de vista de la economía africana la expansión europea fue un elemento exógeno, de la misma forma en que el terremoto de 1923 fue un shock exógeno para la economía japonesa; ambos pueden ser descartados con

seguridad para el propósito del análisis causal, aun cuando exista una teoría que nos dice que fueron inevitables.

Hay algo más que agregar. Cuando se evalúa la importancia causal de un factor exógeno para el desarrollo de un sistema, el factor puede haberse asumido como si no hubiera estado presente, aun cuando desde otro punto de vista hayamos sabido que hubiera sido inevitable. Por cierto que debemos asegurarnos de que la teoría que utilicemos cuando vamos del antecedente al consecuente en sí misma no excluya al antecedente por ser ilegítimo, pero las teorías concernientes a un antecedente *exógeno* no necesariamente tienen que tener un impacto sobre un mundo en el que el antecedente se supone que está ausente. Por otra parte, no podemos suponer la ausencia de factores exógenos si, como en el caso que estamos estudiando, estamos también preocupados con una búsqueda del "mejor pasado". Podemos discutir la causalidad como si lo imposible fuera posible, pero no la moralidad. Más precisamente, la distinción entre teorías concernientes al sistema y teorías concernientes al medio no tiene importancia para los temas de bienestar que entran en juego. Si sostuviéramos que hubiera sido mejor para los africanos si los europeos (*per impossibile*) los hubieran colonizado de una forma completamente benigna o los hubieran dejado completamente tranquilos, entonces podríamos también decir que hubiera sido mejor que los africanos se hubieran rehusado a meterse en el comercio de esclavos o hubieran organizado grupos de estudio para asimilar la tecnología de Occidente. El estándar con respecto al cual se ha de comparar el desarrollo real es el de las opciones factibles *de ambos lados* que habrían llevado al mejor desarrollo de Africa.

Optimismo y pesimismo[34]

Durante 20 años o más, los historiadores británicos han estado envueltos en la búsqueda de *su* mejor pasado, en lo que se llama la controversia del optimismo contra el pesimismo en la interpretación de las consecuencias de la Revolución Industrial en Gran Bretaña. En realidad, no creo que esta terminología sea muy útil. Hay una cantidad de temas distintos en danza, y argumentaré que para clasificarlos es mejor (así como más correcto desde el punto de vista etimológico)[35] hablar de optimismo contra no optimismo y de pesimismo contra no pesimismo. Muy brevemente, hablaré de pesimismo contra no pesimismo como una controversia sobre el desarrollo real: el nivel de vida de los trabajadores británicos durante, digamos, el período 1820-40 ¿subió o bajó? Por el contrario, el debate sobre el optimismo contra el no optimismo concierne al tema contrafáctico de si la tendencia en los salarios reales podría haber sido más alta de lo que realmente fue y, si tal fuera el caso, bajo qué disposicones institucionales. Consideraré tres cuestiones presentadas por Richard Hartwell y Stanley Engerman en un aporte reciente[36] al debate:

1. ¿La clase obrera mejoró su nivel de vida durante la Revolución Industrial?

2. Dado un conjunto de cambios exógenos: ¿la clase obrera estuvo mejor de lo que hubiera estado en ausencia de industrialización?

3. Dada la Revolución Industrial: ¿podría haber habido alguna disposición institucional que hubiera hecho que la clase obrera estuviera mejor de lo que realmente estuvo?

La primera de estas cuestiones concierne al tema del pesimismo tal como lo definimos con anterioridad. Para simplificar interpretaré el problema del nivel de vida como un problema concerniente a la tendencia en los salarios reales, aunque otros numerosos aspectos —tanto cuantitativos como cualitativos— sean pertinentes. Una comparación de dos estados de la economía puede entonces ser llevada adelante de varias maneras distintas, tres de las cuales serán brevemente cubiertas aquí. En primer lugar, podemos utilizar el método "purista" de Samuelson y Sen, discutido en el capítulo 2. La característica peculiar de este enfoque, como se recordará, es que requiere que se lleve a cabo una contrafactibilización aun para la comparación intramundo de los estados reales. El hecho es que una reducción en la satisfacción de una necesidad real puede ser acompañada de un aumento en el potencial para la satisfacción de una necesidad si, por ejemplo, parte de este potencial es desviado a usos militares. Esto, sin embargo, nos lleva a la segunda de las tres cuestiones presentadas por Hartwell y Engerman; retornaremos a este problema más adelante. En segundo lugar, podríamos adoptar el criterio de máxima propuesto por John Rawls en un contexto bastante diferente: de acuerdo con este criterio, los trabajadores de 1840 estaban mejor que los de 1820 si el subconjunto que estaba en peores condiciones del primer grupo estaba mejor que los del subconjunto que estaba menos bien del segundo, tal que "menos bien" fuera definido como "todas las personas con menos de la mitad del ingreso y la riqueza media".[37] Esto puede parecer mal definido, ya que puede haber diferencias significativas dentro de este subconjunto, pero Rawls nos dice que tomemos el promedio del subconjunto como representativo del subconjunto en total. Comparar los ingresos promedio de los *grupos* que están en condiciones no tan buenas en este sentido no es lo mismo que comparar los ingresos de las *personas* que están peor, aun cuando Rawls es a menudo parafraseado en esos términos. Un alto valor para la primera cantidad muy bien puede ir junto con un valor muy bajo para la segunda; por cierto, algunos críticos de Rawls han argumentado que éste es un caso típico en las sociedades modernas.[38] Sin embargo, dejando de lado este problema, no creo que los historiadores estuvieran muy felices con el criterio máxima. Abstraer *completamente* del destino de *todos* los grupos antes mencionados a los que están en peor situación es una simplificación demasiado heroica para el historiador social preocupado por el conjunto de este complejo cuadro.

En tercer lugar, por lo tanto, debemos buscar alguna forma de sopesar los cambios en ingresos de los diferentes grupos de ingreso. En lugar de decir (como Rawls hubiera querido) que una reducción en el ingreso en el grupo que estuviera peor no puede ser compensada por *ningún* aumento, aunque éste fuera de dimensiones astronómicas, en el ingreso de otros grupos, debemos de alguna manera definir un balance [*trade-off*] entre los grupos de distintos ingresos. Va de suyo que este balance no puede ser determinado de manera científica, del mismo modo que no podemos evaluar científicamente la relativa importancia de la equidad y la eficiencia[39] en las sociedades modernas. Esto, sin duda, explica la caldeada naturaleza ideológica de la controversia del nivel de vida.

Para simplificar, podemos suponer que estamos interesados sólo en el ingreso promedio y la dispersión alrededor de la media. Si el ingreso de los que están relativamente peor cuenta más en la función de bienestar social que el ingreso de los que están mejor, entonces tenemos un balance entre promedio y dispersión: una tendencia alcista en los ingresos promedio puede ser compensada por una tendencia hacia una distribución más desigual del ingreso, no sólo entre trabajadores y no trabajadores, sino también, como observó Eric Hobsbawm,[40] entre diferentes estratos dentro de la clase obrera. La distinción entre promedio y dispersión es importante por lo menos en otros dos aspectos. Además del problema de la distribución entre individuos existe, primeramente, el tema de la distribución regional y, en segundo lugar, el tema de la distribución temporal. Con respecto al primero, valores promedio constantes e incluso en baja pueden enmascarar una mejora en el bienestar social causada por mejores medios de transporte, que permiten la rápida transferencia de mercaderías desde regiones con un superávit temporario a regiones con un déficit temporario. Es bastante posible que las inversiones en los ferrocarriles bajen el promedio de consumo (*qua inversión* en ferrocarriles) y suban el promedio de bienestar social (*qua* inversión en *ferrocarriles*). Con respecto al segundo, podemos notar que el balance entre el promedio temporal y la dispersión es más accesible analíticamente que el balance interpersonal, por lo menos en tanto y en cuanto estemos tratando con los mismos individuos en los dos estados. Esto es así porque podemos hacer uso aquí de las preferencias temporales de los individuos mismos: si un individuo vivo tanto en 1820 como en 1840 prefiere las pequeñas fluctuaciones alrededor del promedio bajo en el primer período a las fluctuaciones más grandes alrededor del promedio más alto en el segundo (suponiendo que éste fuera en realidad el esquema), entonces nos podríamos sentir justificados en decir que experimentó un deterioro en su nivel de vida. (Asumiendo, por supuesto, que sus preferencias fueran constantes, lo que puede ser bastante poco realista cuando el individuo se está haciendo viejo y su medio cambia.)

En realidad, la dispersión temporal envuelve dos problemas distintos: la disminución de la utilidad marginal del consumo y la aversión al riesgo. El consumo de $A + a$ en el año 1 y de $A - a$ en el año 2 es, para la mayoría

de la gente una alternativa menos preferida que un consumo igual de $A - e$ en ambos años, para algún e positivo.[41] Así, R. M. Hartwell parece estar justificado al acusar a E. Hobsbawm de "falsedad" cuando éste trata de reducir las consecuencias "optimistas" (o "no pesimistas") de la reducción de la mortalidad señalando que esto puede ser debido a una mayor regularidad de la oferta más que a un aumento del consumo promedio.[42] Si estamos interesados en el bienestar de los individuos más que en el consumo agregado (y la citada observación de Hobsbawm, de que un promedio creciente puede ser compensado por una mayor dispersión, demuestra que, en realidad, ésta es su preocupación), entonces una mayor regularidad de la oferta puede, parcial o completamente, compensar una oferta promedio reducida. Con respecto a la aversión al riesgo, se ha dicho[43] que la mera inseguridad creada por el suministro irregular puede haber reducido el nivel de bienestar de conjunto. Una analogía en el caso interpersonal es la envidia, como demostraré luego.

Se supone a menudo (pero no universalmente[44]) que el bienestar-individual-en-un-momento-determinado es una función exclusivamente del consumo-individual-en-un-momento-determinado. En el párrafo anterior indiqué que la estructura del consumo en el pasado puede entrar en la utilidad[45] presente, y de manera análoga el consumo de otra gente puede tener un impacto sobre mi propio bienestar. Si, por ejemplo, todos los individuos trataran de aumentar (trabajando horas extras y así privándose del tiempo libre) la diferencia entre su propio consumo y el consumo promedio, nadie lo alcanzaría, pero todos tendrían que trabajar más para mantenerse al mismo nivel. Parecería cierto que esta externalización era menos importante durante el período en cuestión de lo que lo es hoy en día, pero sería apresurado concluir que sólo importaban los niveles absolutos de ingresos.

Antes observé que las posiciones políticas de los protagonistas de la controversia estaban reflejadas en su evaluación del nivel de vida. Típicamente, los no pesimistas son conservadores o pensadores libertarios que tienden a dar mayor peso al promedio que a la dispersión alrededor de la media, mientras que los pesimistas son igualitarios, en mayor grado. Además de este matiz, hay también una elección entre bienestar promedio y agregado como criterio de máxima apropiado, dejando de lado todos los problemas de distribución. Los no pesimistas, creo, han puesto el acento en el bienestar total más que en el promedio, lo que favorece por cierto un punto de vista no pesimista en una época de rápido incremento de la población. En una de las más resonantes contribuciones ideológicas al debate, F. A. Hayek escribe:"El proletariado que se puede decir que el capitalismo ha«creado» fue así no una proporción de la población que hubiera existido sin él y que éste había degradado a un nivel más bajo; era una población adicional, a la que se le permitió crecer por medio de las nuevas oportunidades de empleo que le proveyó el capitalismo."[46] Esto creo que debe leerse como una defensa del capitalismo en términos de la maximización del bienestar agregado. Es imposible refutar esto haciendo

referencia al significado usual del término "estándar de vida", que de seguro es independiente del tamaño de la población, ya que la cuestión subyacente real es si la industrialización era de conjunto una cosa buena o mala para la clase obrera. No pretendo tener la solución a la dura pregunta de si muchos trabajadores pobres son mejores (*ceteris paribus*) que pocos trabajadores pobres, excepto para ofrecer el comentario obvio de que desear que los trabajadores que existen no existieran ahora difiere drásticamente de desear que nunca hubieran nacido, y más aún desear antes que hubieran nacido (y antes que cualquiera supiera que ellos serían ellos) que no hubieran nacido. Tomando el punto de vista *ex ante*, creo que puede argumentarse a favor de la medida de bienestar promedio, pero no estoy muy seguro.

Más que explorar otras dimensiones de la controversia sobre el pesimismo, paso ahora a los dos contrafácticos puestos a discusión por Hartwell y Engerman. Con respecto al destino de los trabajadores en ausencia de la industrialización, Hartwell y Engerman muy pertinentemente observan que la noción de factores exógenos debe ser resuelta antes que podamos dar una respuesta. Uno de sus ejemplos se refiere a las guerras napoleónicas, que claramente tuvieron un efecto negativo sobre las condiciones de vida, al desviar parte del potencial para satisfacer necesidades. ¿Estas guerras deberían ser tratadas como exógenas, de tal modo que deberían ser parte del mundo en el que no hubo industrialización? ¿O deberíamos más bien sostener que el capitalismo fue la causa de la industrialización y las guerras napoleónicas (y que no hubo una causa potencial preexistente, que hubiera causado las guerras aun en la ausencia del capitalismo), de tal modo que no podemos eliminar una sin eliminar a la otra? Este ejemplo es puramente hipotético, ya que ningún historiador (por lo que sé) ha mantenido el segundo punto de vista. Mejor ejemplo de una controversia histórica sobre factores exógenos lo provee la caída en la tasa de mortalidad. R. M. Hartwell argumentó que una mortalidad más baja debía ser imputada a la industrialización misma, mientras que H.J. Habakuk, sobre la base de datos más recientes, concluye que fue debido a la vacuna contra la viruela y habría tenido lugar aun si la industrialización no hubiera ocurrido.[47] Esto, de paso, es también pertinente para la cuestión discutida en el párrafo anterior. Con respecto al hecho de que la población creció porque la gente (es decir, la misma gente que hubiera estado por allí de todos modos) vivió más que anteriormente, éste es sin ambigüedades un aumento en bienestar, aun si el nivel de vida (medido por criterios instantáneos, tales como salarios reales) fuera constante, pero, en la medida en que este aumento surgió de la mayor fertilidad, la ambigüedad continúa.

Con respecto a la cuestión de optimismo versus no optimismo presentada por este contrafáctico, tenemos sólo dos alternativas para comparar: el desarrollo real y el desarrollo sin industrialización y sus concomitantes, cualesquiera que éstos sean. Presumiblemente, los optimistas tenderán a clasificar como exógenos los elementos con un impacto adverso sobre el nivel de vida, tales como las guerras, y endógenos los

elementos positivos, tales como la caída en la tasa de mortalidad. Cuando pasamos a la cuestión contrafáctica de si la industrialización hubiera podido tener un curso más parejo en caso de haberse dado cambios adecuados en el conjunto de las instituciones, surgen alternativas y la búsqueda del mejor pasado se hace más compleja. La Revolución Industrial real tuvo lugar en Gran Bretaña, en el contexto de un capitalismo imperfecto: un modo de producción basado en la propiedad privada de los medios de producción, trabajo asalariado, y empresarios que maximizaban la ganancia, pero con un factor de movilidad imperfecto, algunos elementos monopolistas y un gran enclave de industria doméstica y artesanal. Por lo que sé, nadie ha argumentado que este medio real era el mejor posible, de tal modo que los participantes en el debate son todos distintos tipos de no optimistas. Un tipo sostiene que las consecuencias sociales de la industrialización fueron innecesariamente duras debido al contexto capitalista en el que tuvo lugar, mientras que otro sostiene que la miseria se debió al proceso que estaba teniendo lugar en el contexto de un capitalismo *imperfecto*. La segunda posición, a grandes rasgos, tiende a ver al proceso real como mejor de lo que hubiera ocurrido si el capitalismo no hubiera existido y peor de lo que hubiera ocurrido en un contexto de capitalismo perfecto.

Así W.O.Henderson y W.H. Chaloner ponen el acento concretamente,[48] como muchos optimistas lo hacen implícitamente, en que las peores condiciones se encontraban en los sectores de la industria doméstica más que en las fábricas, concluyendo, presumiblemente, que lo bueno en el período en cuestión se debió al capitalismo industrial y lo malo a los sectores que sobrevivían del antiguo régimen. Como correctamente señaló Hobsbawm,[49] esta comparación entre dos sistemas coexistentes en un punto dado de tiempo no puede servir como base para una comparación entre dos sistemas hipotéticos "puros" o sin adulteraciones, ya que un sistema puede tener impacto sobre el otro y aun crear partes del otro. Si los salarios eran bajos en el sector doméstico, esto puede haberse debido al impacto del sector fabril. Más importante aún, el caso de los tejedores de telar a mano puede citarse para demostrar que una vasta clase de trabajadores domésticos puede deber su propia existencia —su prosperidad inicial y subsiguiente miseria— al sistema fabril. Como ya lo observaran Marx y Rosa Luxemburgo, el "capitalismo puro" es una ficción. El modo de producción capitalista ha generado siempre —y más tarde absorbido— a sectores de pequeños artesanos y de servicios que trabajan sobre una base casi artesanal, en los intersticios entre las empresas capitalistas.[50] Así, el meramente asumir esto es ilegítimo, haciendo que la afirmabilidad o la no afirmabilidad del contrafáctico resulte indiferente. El mismo argumento puede ser utilizado con respecto a la idea de Hayek[51] de que el factor de movilidad imperfecta y los elementos monopólicos fueron las "verdaderas razones" del desastre social. Aun si el capitalismo *per impossibile* hubiera nacido completo en forma perfecta, pronto habría desarrollado los elementos monopólicos que, de acuerdo con Hayek, eran un legado del pasado. La

separación que hace Hayek entre el capitalismo y sus imperfecciones es doblemente dudosa: utilizando un criterio dinámico de legitimidad carece de significado y utilizando un criterio estático no es aseverable.

Las Leyes de Navegación

El problema de cuantificar el peso de las Leyes de Navegación surgió en el contexto de determinar las causas de la Revolución Norteamericana, cuando se pensaba que la importancia de los perjuicios económicos entre las causas de la revolución podía ser evaluada midiendo la pérdida económica sufrida por las colonias, como resultado de tener que enviar todas sus exportaciones e importaciones vía Inglaterra. Como un punto previo, podemos observar que este enfoque presupone que es posible especificar el vínculo entre las pérdidas económicas objetivas y los perjuicios sentidos subjetivamente. Este vínculo no puede ser el mismo en todas las situaciones históricas; no puede haber una ley general que establezca que cuando las pérdidas por la expoliación exceden x% del producto bruto nacional ocurrirá una revolución. En primer lugar, la distribución de la pérdida, de manera pareja o despareja, entre la población es crucial para el potencial revolucionario de la pérdida;[52] en segundo lugar, los grupos en cuestión deben estimar la magnitud de la pérdida correctamente; en tercer lugar, la pérdida debe ser adscrita correctamente a la expoliación; en cuarto lugar, deben estar presentes las condiciones para la acción colectiva, como se discute en el capítulo 5. Así, en ausencia de evidencia directa sobre las motivaciones individuales, la cadena de razonamiento debe ir aproximadamente como sigue:

(1) Para el grupo *X* la pérdida por la explotación fue grande.
(2) Para el grupo *X* estuvieron presentes las condiciones estructurales para percibir e interpretar correctamente la pérdida.
(3) Para el grupo *X* estuvieron presentes las condiciones estructurales para la acción colectiva.
(4) Hubo acción colectiva.
(5) Por lo tanto, existe la probabilidad de que el grupo *X*, motivado por la pérdida debida a la explotación, participara en la acción colectiva.

Aquí no cometemos la falacia estructuralista discutida en el capítulo 5, ya que se supone explícitamente que la capacidad para superar el problema de los que actúen por sí solos está a mano. Ahora, el análisis contrafáctico de las Leyes de Navegación no ha ido más allá del primer eslabón de la cadena, o tal vez ni siquiera tan lejos, ya que no se ha hecho ningún esfuerzo para determinar la distribución de la pérdida total en los distintos grupos económicos en las colonias. Si ocurriera que algunas de las condiciones estructurales estuviera ausente, el intento de calcular la pérdida objetiva probablemente hubiera sido una total pérdida de tiempo, un

intento de obtener precisión en el segundo decimal mientras que se ignoraba el primero.

En los primeros estudios sobre el impacto de las Leyes de Navegación, se siguió un enfoque bastante simplista. Supongamos que el precio real por unidad vendida fuera p_1, la pérdida por unidad por enviarlo vía Inglaterra fuera c y el volumen de ventas fuera v. Entonces, la pérdida total para las colonias debida a la explotación inglesa se calculaba como cv, asumiendo que en un mundo contrafáctico sin las Leyes el precio hubiera sido $p_1 - c$ y el volumen de ventas v. Ahora ambas partes de esta presunción son no válidas, la segunda más obviamente que la primera. Asumir ventas constantes con un precio diferente puede ser o solemnemente caracterizado como una presunción implícita de una curva de oferta completamente rígida,[53] o simplemente reconocido como un ejemplo de la falacia preteórica *ceteris-paribus*, discutida en la primera sección anterior. La primera parte de la presunción no es válida por razones más sutiles. En este ejemplo, como en general en mucho del análisis del costo-beneficio, las pérdidas o los beneficios de las situaciones contrafácticas pueden dividirse en dos: el impacto *sobre* las actividades tomadas en la situación real y el impacto *de* las actividades adicionales tomadas en la situación contrafáctica. No sólo es incorrecto discutir pérdidas o beneficios sobre la base sólo de las primeras; puede incluso ser imposible determinar las primeras sin tomar en cuenta los segundos. En el caso presente, la caída en el precio unitario inducida por la eliminación de los costos extra de transporte será menor que estos costos, porque los compradores adicionales subirán los precios. La figura 1 clarifica este punto:

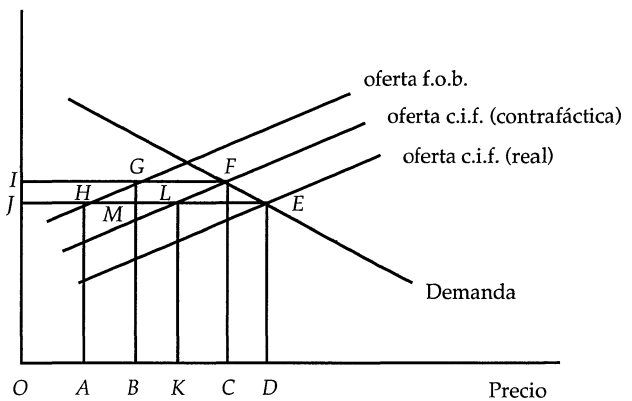


Figura 1

Aquí, la distancia *KD* representa el costo del envío vía Inglaterra. En el mundo real la oferta y la demanda son iguales (en Amsterdam) al precio *OD* y al volumen *DE*, dando un precio f.o.b. en Filadelfia de *OA* y el ingreso total del productor de *OA-DE*. En la situación contrafáctica, la oferta y la demanda son iguales al precio c.i.f. OC con volumen *CF*, dando un precio f.o.b. de *OB* y un ingreso total del productor *OB–CF*. Es fácil ver que la diferencia en ingreso (es decir la diferencia entre los rectángulos *OAHJ* y *OBGI*) no tiene nada que ver con las ventas reales multiplicadas por los costos de transporte (rectángulo *KDEL*), ni con la magnitud precedente más las ventas extra multiplicadas por el precio real más los costos de transporte, ni con cualquier otra magnitud concebible. Si la diferencia estimada en el ingreso de este caso coincidiera con la medida simplista *KDEL*,[54] esto no puede ser más que un accidente. Es muy posible trazar las curvas de tal manera que la medida simplista no pueda siquiera servir como una estimación mínima de la pérdida. En muchos contextos, el impacto de la presunción contrafáctica sobre las actividades existentes puede servir como una estimación mínima, pero no si los efectos de la interacción están presentes. Tomemos el ejemplo de una nueva carretera, que al reducir la distancia ahorrará dinero (es decir, gasolina) para los conductores que existen y creará beneficios para los conductores adicionales. Si consideramos solamente el costo monetario, los ahorros creados para los conductores existentes pueden servir como una estimación mínima de los ahorros totales. Si, por el contrario, los conductores adicionales atraídos por la carretera generaran un costo extra en *tiempo* para los conductores existentes al aumentar la congestión en la ruta, los efectos de la interacción están presentes y podría concebirse que fuera el caso que el efecto de molestia de los nuevos conductores sobre los viejos superara el beneficio para los nuevos, por lo que la medida simplista no serviría como estimación mínima.

Esto también se puede demostrar en términos de análisis de equilibrio parcial contra análisis de equilibrio general. Lo que uno hace en el enfoque simplista es asumir que una sola unidad fue enviada directamente desde las colonias norteamericanas a su destino en Europa, mientras que *todas las demás unidades fueron enviadas vía Inglaterra*. En este análisis parcial tomamos el precio como dado y luego calculamos el beneficio que le resultaría al único productor que podría actuar de esta manera. Al generalizar el resultado del análisis de equilibrio parcial uno llega entonces a una conclusión obviamente incorrecta sobre el equilibrio general, que surgiría si todos pudieran enviar directamente a destino. El productor marginal no puede influenciar los precios; cada productor puede verse como el productor marginal; por lo tanto, las conclusiones válidas para el productor marginal son válidas para todos los productores tomados simultáneamente. Esto, por supuesto, una vez más, es una mera falacia de composición. Acepto que en muchos análisis contrafácticos esta falacia funciona, cuando las contrafactibilizaciones agregadas son estimadas mul-

tiplicando el im-pacto de un cambio contrafáctico sobre una sola unidad por la cantidad de unidades.

Otro error relacionado con el anterior se puede encontrar en el intento de R.P.Thomas de determinar las ventas y los precios en el contrafáctico 1770. Escribe[55] que, "El precio que hubiera existido en las colonias en ausencia de las Leyes de Navegación puede ser estimado...dividiendo el precio observado en Amsterdam para el tabaco de Virginia antes de la revolución por la proporción de los precios del tabaco de Amsterdam a Filadelfia después de la Revolución". En otras palabras, el precio contrafáctico f.o.b. 1770 es el precio real c.i.f. 1770 multiplicado por

$$\frac{\text{precio real f.o.b. 1790}}{\text{precio real c.i.f. 1790}}$$

Este es un intento de sacar conejos contrafácticos de sombreros reales; aun si aceptáramos la dudosa noción de que la proporción de los precios f.o.b. a c.i.f. hubiera sido la misma en el contrafáctico 1770 como en el 1790 real, el precio contrafáctico f.o.b. sólo puede ser obtenido si sabemos el precio *contrafáctico* c.i.f. de 1770. El procedimiento propuesto por Thomas significa suponer que el precio contrafáctico c.i.f. de 1770 hubiera sido el mismo que el precio c.i.f. real de 1770, lo que es equivalente a asumir una curva de demanda completamente elástica, una presunción que es no menos arbitraria que asumir una curva de oferta completamente rígida.

Los ferrocarriles y el crecimiento económico norteamericano

Este es el título de un libro escrito por Robert Fogel, que constituye por cierto el intento más ambicioso de contrafactibilización en grande jamás llevado adelante por un historiador serio.[56] Suponiendo que el ferrocarril no existiera, Fogel se pregunta si esto hubiera provocado una diferencia sustancial en el PBN norteamericano de 1890. En su intento de contestar esta pregunta, Fogel explícitamente elige el método de insertar una economía sin ferrocarril en el pasado real, yendo hacia atrás en el tiempo hasta el momento (alrededor de 1830) cuando uno podría legítimamente asumir un punto de ramificación sin ferrocarriles. Por lo menos al nivel de las intenciones, sigue el procedimiento propuesto con anterioridad. Veremos, sin embargo, que es dudoso si en realidad saca todas las implicancias que se siguen de su elección. En algunos aspectos su procedimiento es más cercano a lo que presumiblemente Lewis hubiera recomendado: tomemos la economía real de 1890, quitemos los ferrocarriles y asumamos un mínimo de otros cambios, y calculemos otra vez el PBN. En

otros aspectos toma el método muy seriamente discutiendo en asombroso detalle las posibilidades alternativas al ferrocarril que hubieran aparecido en el período que lleva hasta 1890.

Como de costumbre comenzamos con la cuestión de la legitimidad, porque, si ésta no puede ser contestada afirmativamente, no surge el tema de la afirmabilidad. ¿Es legítimo entonces suponer que el ferrocarril no fue nunca inventado? La respuesta a esta pregunta depende de cuál de las teorías del cambio tecnológico que están en estos momentos discutiéndose preferimos: una centrada en la oferta o una centrada en la demanda; una teoría que ponga el acento en la inevitabilidad de las invenciones o una teoría de las invenciones como algo esencialmente azaroso, como las mutaciones biológicas. En las teorías no deterministas, todos (o, en teorías moderadamente no deterministas, muchos) contrafácticos relacionados con la no invención son legítimos. Bastante sorprendentemente veremos que ambas variedades del enfoque determinista también permiten algún espacio para la remoción mental de los inventos realmente realizados.

En la teoría determinista de la demanda de Jakob Schmookler[57] un crecimiento en la industria induce un inevitable y predecible aumento en el esfuerzo inventivo dirigido hacia la industria y en la cantidad de invenciones realmente hechas para esa industria. Este determinismo, sin embargo, concierne principalmente al extremo receptor y no al extremo productor de la invención. Podemos predecir que cuando la industria ferroviaria crece, habrá un aumento en la cantidad de patentes para una mejor tecnología ferroviaria, pero no podemos decir mucho con respecto a las técnicas utilizadas en estas invenciones, es decir si sus principios serán mecánicos, eléctricos o termodinámicos. Esta ignorancia con respecto a los detalles materiales y técnicos de los inventos no es una simple cuestión de la imposibilidad lógica de predecir en detalle una invención futura (si pudiéramos predecirla, la habríamos inventado). El punto de vista de Schmookler es más bien que los detalles son objetivamente azarosos, en el sentido de que no se nos permite formular ninguna afirmación contrafáctica de que si el inventor no hubiera realizado su invento, entonces alguien más lo hubiera hecho. Todo lo que podemos decir es que alguna invención funcionalmente equivalente probablemente se hubiera encontrado. Si estamos principalmente interesados en los aspectos funcionales del invento, ésta no es una restricción seria. Para algunos propósitos, sin embargo, estamos igualmente preocupados por el sustrato material del invento. Cuando Fogel intenta determinar la importancia para la economía norteamericana de la demanda de hierro en lingotes inducida por el ferrocarril, estamos tratando con un aspecto material del ferrocarril, y no con uno funcional. Si suponemos que en ausencia del ferrocarril, el motor de combustión interna y toda la industria automotriz se hubieran desarrollado unos 50 años antes de lo que ocurrió, esta es una alternativa que en algunos aspectos es funcionalmente equivalente (hace muchas de las mismas cosas), pero materialmente muy diferente (implica una demanda

más pequeña para el acero y más grande para el caucho). Por supuesto, la teoría de Schmookler no nos permite suponer que el motor de combustión interna hubiera aparecido en la escena tecnológica, sino sólo que *algún* equivalente funcional con requerimientos materiales probablemente diferentes lo hubiera hecho.[58]

En una teoría determinista de la oferta, tal como la propuesta por Robert Merton y otros estudiosos de las invenciones simultáneas,[59] no sólo los aspectos funcionales sino también los aspectos materiales de un invento son inevitables, en el sentido de permitir afirmaciones contrafácticas, al efecto de que si su inventor no, algún otro lo hubiera hecho. A lo sumo se permite un retraso en el tiempo, y el rol del genio sería hacer que los descubrimientos se hicieran antes de lo que hubiera sido hecho de otra manera. Este tipo de teoría puede parecer excluir todas las presunciones del tipo de la no existencia del ferrocarril, pero si uno mira más atentamente las cuestiones no son tan simples. Por ejemplo, si el progreso de la ciencia y la tecnología se ve como un proceso mayormente autónomo, poco o nada influenciado por el medio social y económico (que sí es influenciado por el mismo), entonces podemos permitirnos suponer cosas como la no existencia del ferrocarril, al igual que podemos aceptar aquellas del tipo de que no hubo un terremoto, que antes mencionamos. Sólo si la ciencia, la tecnología, la economía y la sociedad forman un sistema de variables interrelacionadas, ningún sistema de los cuales puede ser separado y resuelto independientemente del resto, debemos imponer severas restricciones a los contrafácticos con respecto al cambio tecnológico y su impacto sobre el desarrollo económico. Por supuesto, la noción de un proceso internamente determinado es bastante poco plausible, y sólo un poco menos para el progreso de la ciencia. El concepto tecnológico de una reacción en cadena (o de una acumulación combinatoria de ideas) no debería ser usado para caracterizar el proceso del cambio tecnológico mismo.

Hay también un tercer tipo de teoría determinista que probablemente atraerá más a la mayoría de la gente que cualquiera de las dos anteriores; una teoría mixta de la oferta-demanda que dice que el efecto conjunto de las condiciones de la demanda y del estado de las artes fue hacer inevitable el surgimiento de los inventos que realmente surgieron.[60] Esta teoría tiene la ventaja de permitir una evaluación de las tecnologías de transporte alternativas que se hubieran desarrollado en una economía sin ferrocarriles, pero sólo al costo de hacer ilegítima la presunción misma de un mundo sin ferrocarriles. Como se afirma anteriormente, si se pudiera contestar la pregunta sobre los sustitutos, entonces no se la debería haber hecho para nada. En realidad, es mucho mejor para Fogel conceder que no se hubieran desarrollado sustitutos e intentar probar que aún entonces la economía sin ferrocarril hubiera tenido en 1890 un PBN no mucho más pequeño que el real (que es la tesis a favor de la que está argumentando), antes que tratar de probar la misma tesis argumentando que los sustitutos que se hubieran desarrollado habrían sido a grandes rasgos equivalentes al propio ferro-

carril. La primera línea argumental (que es la que se sigue en el libro) es perfectamente consistente. No implica que Fogel tenga una creencia positiva de que ninguna de las alternativas se hubiera desarrollado (tal creencia requeriría una teoría, que podría entrar en conflicto con la presunción de la no existencia de ferrocarriles), sino sólo que él está dispuesto a aceptar la presunción de que de cualquier modo está prejuiciado contra la conclusión que quiere probar. Si intentara fortalecer su conclusión por medio de la segunda línea de argumentación, podría serruchar la rama sobre la que está sentado. A menudo, en este tipo de ejercicio se da el caso de que más es menos y que la ignorancia es fuerza.

Suponiendo ahora que la cuestión de la legitimidad ha sido contestada afirmativamente, debemos examinar si le sigue la conclusión sobre la falta de importancia de los ferrocarriles: si el producto bruto nacional de una economía sin ferrocarriles hubiera sido aproximadamente de la misma magnitud que el real. La argumentación a grandes rasgos es la siguiente. Fogel primero subdivide la contribución total de los ferrocarriles al crecimiento de la economía norteamericana en su contribución qua proveedores de servicios de transportes (relacionada con el aspecto funcional) y su contribución *qua* usuarios de varios insumos, especialmente lingotes de hierro (aspecto material). Como Fogel dice haber demostrado (y no tengo nada contra eso) que esta última contribución fue insignificante, la contribución total puede ser igualada al impacto del lado de la provisión. Fogel trata exclusivamente el transporte de *mercaderías*, dejando sin explorar el de pasajeros. Con la importancia que se le da en la reciente teoría económica a los costos transaccionales, esta negligencia es difícil de justificar. Las discusiones cara a cara se facilitaron con el ferrocarril, los contratos se pudieron concluir más rápidamente, los inventarios se ajustaron más rápidamente, etcétera. Volveré a algunos aspectos de este problema, pero primero demos una mirada al tratamiento del transporte de mercaderías de Fogel. Dentro de la categoría general de mercancías, en realidad sólo considera las agrícolas, y dentro de esta subcategoría se restringe aún más al maíz, el trigo, el puerco y la carne vacuna. Aceptaré su argumento de que si su modo de razonamiento es válido para estas mercaderías es válido para todas las mercaderías.

Una distinción crucial en la argumentación es la del transporte interregional (desde los productores en el Oeste Medio a los consumidores en el Este y el Sur) y el transporte intrarregional (de los productores a los centros de distribución y desde los centros de distribución a los consumidores). Para el transporte interregional encuentra que los ahorros sociales generados por el ferrocarril fueron de alrededor del 0,6% del PBN en 1890 (esta cifra, como todas las demás, es propuesta como el tope). Para el transporte intrarregional encuentra que los ahorros fueron de alrededor del 1,2%, de tal modo que el total de las cuatro mercaderías agrícolas sería de alrededor del 1,8%. La generalización a todos los productos entonces da un total muy por debajo del 5% del PBN. Los principales ajustes contrafácticos hechos por Fogel para ir del antecedente al consecuente son los

siguientes. En primer lugar, supone que una gran cantidad de depósitos adicionales serían construidos para asegurar el suministro de carne en los mercados secundarios (Este y Sur) durante los cinco meses del año, cuando los canales no podrían ser utilizados debido al congelamiento. En segundo lugar, supone que el esquema espacial de la producción agrícola habría tenido que ser modificado, para excluir emplazamientos demasiado alejados de las facilidades de transporte en un mundo sin ferrocarril. Tercero, supone que una cantidad de canales nuevos habrían tenido que construirse, y las rutas, mejorarse mucho; en ambos casos se supone que las tecnologías reales fueron las que prevalecieron.

El principal defecto del análisis de Fogel fue muy bien señalado por Paul David:[61] es la confusión de las *ganancias por su introducción* con las *pérdidas por su retiro*. Perder algo es muy diferente a no haberlo tenido nunca; la pérdida puede ser mejor (como en el caso presente) o peor (como en la pérdida de un niño). En la explicación de los contrafácticos de Lewis, las nociones de similitud y de distancia invocan la idea de la pérdida por retiro, de tal modo que el mundo más parecido posible es aquel en el que esta pérdida es minimizada, sujeta a la condición del antecedente. En mi explicación, sin embargo, la idea es que la ganancia de la introducción perdida es la más relevante. Cuando se discute la importancia del ferrocarril para el producto bruto nacional de 1890 no queremos saber qué hubiera ocurrido, si, digamos, repentinamente la fatiga estructural de las vías hiciera este modo de transporte demasiado peligroso para el uso continuado, y todas las demás cosas siguieran iguales. Queremos saber cuáles características de la economía de 1890 hubieran existido si nunca hubiera habido ferrocarriles.

Se debería repetir que la *intención* de Fogel es contestar precisamente esta pregunta, pero —como convincentemente sostiene David— algunos de sus argumentos realmente tratan de la pérdida instantánea debida al retiro del ferrocarril y no de las pérdidas acumulativas debidas a la no introducción del ferrocarril. Un ejemplo —no mencionado por David— se refiere a las técnicas de refrigeración. En 1890 hubiera sido perfectamente posible tener grandes cantidades de carne depositadas por largos períodos, pero antes de 1875 la técnica necesaria para la producción mecánica del hielo no existía. Argumentando de acuerdo con Schmookler, uno podría sostener que en un mundo sin ferrocarril la demanda de instalaciones de almacenaje hubiera llevado a que la máquina de amoníaco fuera inventada antes, o por lo menos a alguna instalación funcionalmente equivalente, pero si preferimos (como creo que deberíamos hacerlo) una teoría de las invenciones mixta (oferta-demanda), esta línea de argumentación puede estar cerrada para nosotros. Deberíamos entonces tener que aceptar que por un largo período el consumo de carne en los mercados secundarios hubiera estado restringido a la mitad estival del año, cuando el ganado vivo podía ser embarcado por agua. No debería ser necesario abundar sobre las consecuencias de esto para el desarrollo de la agricultura

norteamericana,consecuencias que todavía se hacían sentir en el hipotético 1890.

El principal ejemplo de David concierne a la existencia de irreversibles economías de escala en la producción debido al aprendizaje a través de la práctica. Si los menores costos de transporte del ferrocarril llevaran a precios más bajos —y una mayor demanda— de las mercaderías transportadas, se podrían haber desarrollado técnicas de producción en masa que— en el caso del retiro instantáneo de los ferrocarriles en 1890— hubieran tenido todavía una mayor eficiencia ante el volumen de producción *más pequeño* que el de las técnicas inicialmente dadas. Si uno entra y mide la importancia de la escalera en el punto en el que se ha hecho posible patearla afuera, no debería sorprenderse si resulta ser prescindible. Un caso plausible puede presentarse a favor de la proposición de que no resultaría mayor daño (y sí, probablemente, algún beneficio) si hoy en día abandonáramos el sistema de la línea de montaje y retornáramos a los métodos que dan un mayor espacio para las decisiones individuales, pero de esto uno no debiera concluir que la economía habría sido bastante parecida hoy en día si la línea de montaje no hubiera sido introducida.

Haré un último comentario sobre lo que considero la presunción implícita de Fogel, de que el curso de la economía hipotética hubiera sido una secuencia de estados en equilibrio. Una economía real nunca está en equilibro, pero durante largos lapsos puede operar cerca del equilibrio si los mecanismos que hacen a la estabilidad son lo suficientemente sensibles y rápidos. Me parece evidente que la adaptación a los inconvenientes hubiera sido más lento en una economía de canales que en una de ferrocarriles. Las enormes cantidades de comida que hubieran tenido que ser almacenadas para el invierno hubieran requerido ciclos de inventario extremadamente largos, con la correspondiente pérdida de eficiencia. Los cuellos de botella en las relaciones interindustriales hubieran constituido un problema mucho más serio de lo que en realidad fueron; el aumento de los costos de transacción implicados por la ausencia de un transporte de pasajeros rápido ya han sido mencionados. El punto es que la importancia del ferrocarril fue mucho más amplia que la mera reducción de los costos de transporte: al permitir ajustes y comunicaciones casi instantáneos, flexibilidad y previsibilidad, el ferrocarril fue parte de las condiciones generales para el desarrollo económico. Si hubiera habido una previsión perfecta y una certeza total asociada con todas las decisiones económicas, esto no hubiera importado demasiado, pero en un mundo donde perpetuamente se generan consecuencias no queridas, es de la mayor importancia que estas fluctuaciones sean rápidamente frenadas y que no se les permita propagarse, como ocurriría en el prolongado tiempo de reacción de una economía de canal. La teoría dual del cambio social, propuesta en el capítulo 5, incluía buenas instalaciones de comunicación entre las condiciones que facilitaban la acción colectiva con el propósito de restablecer el equilibrio, y es bastante obvio que son también decisivas para la eficacia de los meca-

nismos que restablecen el equilibrio a través de las acciones individuales en el mercado.

Los esclavos norteamericanos y su historia

Concluyo el presente capítulo, y el libro, con una exposición de algunas afirmaciones contrafácticas que han sido propuestas en el análisis de la esclavitud de los negros norteamericanos. Esta discusión ha sido recientemente llevada a un nuevo pico debido a la publicación del libro de Robert Fogel y Stanley Engerman *Time on the Cross*,[62] que se presenta como una revisión drástica de virtualmente todos los principios del saber convencional sobre la esclavitud del negro norteamericano. ¿La esclavitud fue provechosa? ¿Los dueños de esclavos eran racionales? ¿Los esclavos eran eficientes? El saber convencional dice No; Fogel y Engerman, Sí. ¿Se retardó el crecimiento económico del Sur debido a la esclavitud? ¿Los esclavos eran tratados cruelmente? El saber convencional dice Sí; Fogel y Engerman, No. Tal es de todas maneras la imagen de sí mismos de los autores, aun cuando se ha demostrado[63] que nadie mantuvo simultáneamente todos los puntos de vista atribuidos a la sabiduría convencional y que muchas de estas opiniones fueron abandonadas hace mucho tiempo por la mayoría de los estudiosos en este campo. La discusión historiográfica, sin embargo, es sólo una y probablemente la menos importante de las cuatro dimensiones principales del furioso debate que se ha mantenido alrededor de *Time on the Cross*. Una segunda dimensión se refiere a la cuestión de si Fogel y Engerman tienen datos ciertos, sus ejemplos son representativos y su compresión de las fuentes, correcta. Una tercera dimensión se relaciona con la lógica de su inferencia de los hechos o supuestos hechos, tal como su uso de datos agregados como fundamento para las inferencias sobre las motivaciones. La cuarta dimensión es la ética, donde los críticos han argumentado que el "violento" antirracismo de los autores es tan sólo una variedad más sutil de lo que ellos están atacando y que el efecto, si bien no querido, de su trabajo es obviar las distinciones fundamentales entre la libertad y la sujeción.

En lo que sigue, trataré con un pequeño subconjunto de problemas dentro de la tercera dimensión, y no exclusivamente con el trabajo de Fogel y Engerman. Este subconjunto se relaciona con el uso del análisis contrafáctico para contestar dos preguntas distintas. En primer lugar, podemos preguntar cómo habría sido la economía sureña en 1860 si no hubiera habido esclavitud; en segundo lugar, podemos suponer que la esclavitud nunca fue abolida y preguntarnos qué habría ocurrido. La preocupación por la relación entre la esclavitud y el crecimiento económico subyace en ambas preguntas: ¿la esclavitud en realidad impidió el crecimiento económico del Sur, o habría esto ocurrido si se le hubiera permitido continuar? Una respuesta afirmativa a la primera pregunta implica que la esclavitud estaba estancada, mientras que la misma respuesta para la segunda implica

que estaba *moribunda*. Es perfectamente posible afirmar que una economía ha crecido vigorosamente hasta un tiempo *t* y que no hubiera continuado creciendo; éste es por cierto el argumento de la escuela de pensamiento que ve el uso de recursos no renovables como la condición esencial para el crecimiento económico del Sur, antes de la guerra.

Fogel y Engerman responden negativamente a ambas preguntas. Además también presentan un tercer problema, con respecto a las expectativas de los dueños de esclavos al inicio de la Guerra Civil. Ellos sostienen que los dueños de esclavos eran optimistas y que su optimismo era justificado, en el sentido de que si la guerra no hubiera tenido lugar sus expectativas habrían sido confirmadas. Desde un punto de vista metodológico, este procedimiento parece sensato, y por cierto indispensable, si tenemos que evaluar completamente la tesis de una economía sureña moribunda. Supongo que hubiéramos tenido que aceptar esta tesis si se pudiera probar o que los dueños de esclavos mantenían un punto de vista pesimista del futuro o que su futuro hubiera sido oscuro; conversamente, la tesis puede ser desaprobada sólo refutando ambas proposiciones. Si, por ejemplo, el futuro hubiera sido color de rosa, pero por razones que no podrían haber sido previstas por los dueños de esclavos en 1860, tales como un cambio exógeno en la demanda de algodón, el sistema estaba moribundo en un sentido *ex ante*, y los dueños de esclavos hubieran actuado como si estuviera moribundo. La comparación entre el resultado esperado y el resultado contrafáctico es probablemente el pico de la abstracción en el razonamiento histórico. A menudo comparamos los resultados anticipados con los reales y, a menudo por lo menos, el resultado real con el contrafáctico, pero no sé de ningún otro caso donde haya habido ocasiones y por cierto la necesitad de comparar el resultado anticipado sobre la base de asumir ciertos resultados que habrían tenido lugar si estas presunciones hubieran resultado ser correctas. No nos concierne la "capacidad de la clase de los dueños de esclavos para leer la bola de cristal",[64] ya que sabemos que la Guerra Civil probó que estaban equivocados. Más bien queremos saber si hubieran tenido razón si la guerra no se hubiera producido.

Antes de entrar en detalles con esta comparación, sin embargo, tomaré el primer contrafáctico antes mencionado, con respecto al estado hipotético de la economía sureña digamos, en, 1860, en ausencia de la esclavitud. Es extremadamente difícil determinar cuál de los mundos sin esclavitud es el más parecido al real, pero he aquí varios de los candidatos más plausibles:

(1) Tomemos la economía real de 1860; reemplacemos a los dueños de esclavos por grandes propietarios agrícolas y a los esclavos por aparceros y jornaleros; eliminemos todas las características de la sociedad que surgen *directamente* de la institución presente de la esclavitud.

(2) Tomemos la economía real de 1860; hagamos las mismas sustituciones que en el caso precedente; eliminemos todas las características que directa o *indirectamente* surgen de la institución presente de la esclavitud.

(3) Tomemos la economía real de 1860; hagamos las mismas susti-

tuciones que en los casos precedentes; eliminemos todas las características que directa o indirectamente surgen de la presente o *pasada* institución de la esclavitud.

(4) Retrocedamos en el tiempo hasta el último punto de ramificación en el que la esclavitud podría haber sido abolida y desarrollemos una economía alternativa a partir de ese punto hasta 1860.

(5) Retrocedamos en el tiempo hasta el último punto de ramificación a partir del cual el desarrollo podría haber comenzado en el que la esclavitud nunca se hubiera establecido, y desarrollemos una economía alternativa desde ese punto hasta 1860.

De todas estas posibilidades las dos primeras invocan solamente el criterio estático de la legitimidad y son más o menos concordantes con la teoría de David Lewis. Las dos últimas, por el contrario, corresponden al enfoque sostenido por el presente autor. La tercera es un poco ambigua, pero es fundamentalmente una variante de las dos primeras. Por otra parte, la diferencia entre (1), (2) y (4) por un lado, y (3) y (5) por el otro, es que las primeras permiten que la economía contrafáctica retenga todas las características (racismo, etc.) que se siguen de una historia pasada de esclavitud, mientras que las segundas definen a una economía sin esclavitud de una manera más rigurosa, como una sin "rastros de histéresis"[65] de un pasado esclavista. La diferencia entre (1) y (2) es la diferencia entre un estado de no equilibrio y uno de equilibrio. Cuando quitamos la esclavitud debemos también quitar los aspectos de la sociedad que están inmediata y conceptualmente ligados a esa institución, tales como la policía que seguía los rastros de los esclavos fugitivos, las caballerizas para esclavos, el estatus legal diferente para los esclavos, etcétera. Es otra cuestión que quitemos o no también los aspectos que (en equilibrio) están *causalmente* ligados a la esclavitud. Una sociedad sin esclavitud, pero con las características políticas e ideológicas de una sociedad esclavista, es bastante posible, pero no estaría en equilibrio. La transición de la esclavitud a la falta de la misma presumiblemente también requeriría algunos cambios en la naturaleza y la ubicación de las actividades económicas, debido al factor movilidad repentinamente creado y a los cambios en motivación. La distinción entre (1) y (2) corresponde, *mutatis mutandis*, a una distinción entre el Sur real de 1865 y el Sur real de 1880 o alguna otra fecha, en la que se pueda suponer que un equilibrio aproximado sin esclavitud se habría alcanzado.

Por lo que puedo comprender o adivinar, el mundo de la alternativa (1) debe ser el más parecido a un mundo sin esclavitud si aceptamos la explicación de Lewis, pero este mundo es extremadamente carente de interés si estamos preguntando sobre la importancia de la esclavitud para el crecimiento económico del Sur. Si alguien pidiera una prueba de la presunta ineficacia del socialismo, no se sentiría muy impresionado si le señaláramos los problemas que aquejan a una economía inmediatamente después de una revolución socialista violenta. Para evaluar la efectividad de

un orden institucional debemos entrar en el punto en el tiempo —sea real o contrafáctico— donde el sistema haya operado durante un período de tiempo lo suficientemente largo para que las consecuencias de este ordenamiento se hayan producido. No quiero decir que la noción de Lewis del mundo *más próximo* sin esclavitud coincide con el estado que inmediatamente *sucedió* a la esclavitud (o la habría sucedido si ésta hubiera sido abolida). En la teoría de Lewis no se exige que haya un camino posible desde el mundo real hasta el mundo posible más próximo. Sin embargo, no puedo ver ninguna objeción a la interpretación de que el mundo más próximo posible a un estado de equilibrio pueda ser un estado no equilibrado, tal como el estado que inmediatamente sucede a un equilibrio interrumpido es también falto de equilibrio.

Podemos ahora proceder a vincular este análisis algo árido con algunos argumentos contrafácticos sobre el Viejo Sur. Un ejemplo de un modo no poco común de razonar puede tomarse de un artículo de Robert Russel, en el que trata de refutar la tesis de que la esclavitud era dañina para los blancos que tenían esclavos:

> En los tiempos de la esclavitud, las ciudades y los pueblos del Sur, al no ser numerosos ni grandes, derivaban su apoyo principal de los distritos de las plantaciones, en los que había muchos esclavos, y no de las regiones de las pequeñas granjas, en las que había pocos esclavos. Eran principalmente los plantadores los que compraban, vendían, pedían prestado, viajaban y enviaban a sus hijos a las academias y universidades. Parece bastante cierto, por lo tanto, que si no hubiera sido por las plantaciones y la esclavitud, las ciudades y los pueblos del Sur hubieran sido menos y más pequeños, resultando en aun menos oportunidades para los blancos que no tenían esclavos.[66]

El "por lo tanto" de Russel, vinculando la parte fáctica del párrafo con la parte contrafáctica, nos quita la respiración. En primer lugar, ¿cómo podía uno en una economía no esclavista distinguir entre los blancos que tenían esclavos y los que no? ¿Cómo se podía identificar a las personas contrafácticas que en el mundo real eran respectivamente una y la otra? Una posibilidad es buscar a las personas que eran los equivalentes funcionales, es decir, a los que no poseían trabajadores y a los que poseían trabajadores, respectivamente, y ver si el destino de los segundos hubiera sido mejor o peor que el de los blancos que en realidad no tenían esclavos. De acuerdo con la alternativa (1), antes mencionada, esto es plausible, ya que en ese caso uno sabe que está tratando con los mismos individuos en roles diferentes. En las alternativas (2)-(5), sin embargo, no hay seguridad de que ése sea el caso, y entonces cualquier afirmación sobre la distribución contrafáctica de los ingresos no implica nada sobre el destino de los que no eran dueños de esclavos. Como argumenté que la alternativa (1) es en sí misma no plausible, esta cuestión no tiene una respuesta satisfactoria. (Otra posibilidad, presuponiendo las alternativas (4) o (5), sería identificar

a las personas contrafácticas con personas reales mirando a sus antepasados comunes antes del punto de ramificación, pero en ese caso existe la dificultad de que muchos de los que en realidad no son dueños de esclavos podrían no haber nacido nunca en el mundo contrafáctico. En segundo lugar, recibimos la clara impresión de que, de acuerdo con Russel, los dueños de esclavos simplemente hubieran desaparecido sin ningún sustituto en el mundo contrafáctico, ya que parece eliminar la idea de los homólogos contrafácticos para los plantadores, tales como los grandes granjeros capitalistas, que podrían haber sido funcionalmente equivalentes para proveer apoyo a las ciudades y pueblos.

Así, el mismo *contresens* está presente en un muy conocido artículo de Robert Fogel, "The specification problem in economic history". En el curso de una discusión de la afirmación de Eugene Genovese de que el desarrollo industrial del Viejo Sur se vio retrasado por la esclavitud, Fogel escribe que

> la afirmación de que el tamaño del mercado sureño de mercancías estaba limitado por la existencia de la esclavitud implica la proposición contrafáctica de que en ausencia de la esclavitud el mercado sureño de mercancías manufacturadas habría sido mayor de lo que en realidad fue. Dados los grupos de consumo en los que Genovese divide la población, la condición básica para que esta proposición sea válida es
>
> $$(D'_0 - \bar{D}'_0) + (D''_0 - \bar{D}''_0) + (D_s - \bar{D}_s) + (D_f - \bar{D}_f) > 0,$$
>
> donde
> D'_0 = función de demanda de los dueños de esclavos de mercancías manufacturadas que serían utilizadas en el mantenimiento de esclavos;
> D''_0 = función de demanda de los dueños de esclavos de mercancías manufacturadas para todos los demás usos;
> D_s = función de demanda de los esclavos de mercancías manufacturadas;
> D_f = función de demanda de mercancías manufacturadas de las personas libres que no tenían esclavos.
> Los símbolos con barras refieren al período esclavista; aquellos sin barras refieren al período abolicionista.
> Obviamente la condición básica podría haber sido satisfecha de una multitud de maneras. Si sólo se permitiera la comparación ordinal de las funciones por pares, hay 240 maneras de satisfacer la desigualdad. Genovese no especificó claramente cuál de estas alternativas era la que tenía en mente. Lo más que puede razonablemente inferirse de este análisis es que $D_s - \bar{D}_s > 0$ y que $D_f - \bar{D}_f > 0$. Las restricciones reducen las alternativas teóricas, pero la cifra que queda, sesenta, es todavía muy grande.[67]

Con enemigos como éste, Genovese no necesita amigos. Fogel está tan deseoso de burlarse de la falta de sofisticación de Genovese que llega a invocar construcciones teóricas desprovistas de significado. La noción D'_0, por ejemplo, no tiene ningún significado. (O tal vez uno podría sostener que debería ser exactamente igual a cero, lo que entonces reduciría aún más la cantidad de alternativas.) Más importante aún, la

distinción entre D''_0 y D_f es imposible de establecer, por las razones antes dadas. Fogel curiosamente se refiere a la situación contrafáctica como "el período abolicionista", pero por supuesto el período real posterior a 1865 no puede servir como un sustituto del contrafáctico 1860; tampoco es obvio que este contrafáctico 1860 deba ser un mundo en el que la esclavitud haya sido *abolida* y no un mundo en el que ésta hubiera desaparecido en silencio o nunca hubiera sido introducida.

Las afirmaciones hechas aquí pueden generalizarse de la siguiente manera: bajo todas las condiciones plausibles es imposible evaluar el impacto de un sistema económico sobre grupos que son definidos en términos de nociones específicas de ese sistema, ya que estos grupos no pueden ser identificados en mundos en los que el sistema se supone que está ausente. Podemos indagar en el destino de los pequeños productores agrícolas o los dueños de trabajadores en un Sur capitalista, pero no sobre lo que hubiera pasado a los libres que no tenían esclavos. El intento de probar que los segundos fueron víctimas de la esclavitud es tan carente de significado como el intento de Russel de probar lo contrario. Asimismo, no podemos determinar el impacto del capitalismo sobre los trabajadores, mientras que *sí* es posible analizar el impacto de la agricultura capitalista sobre las formas precapitalistas que coexistieron con ella o el impacto del capitalismo temprano sobre los sectores artesanos, porque en los segundos es posible eliminar el medio sin quitar también las condiciones para identificar a los grupos en cuestión. El problema puede ligarse al análisis de Sen del ingreso nacional, como lo disentí brevemente en el capítulo 2. El único caso donde podemos decir sin ambigüedades que los libres sin esclavos hubieran estado (digamos) en mejor situación sin la esclavitud, es cuando se puede demostrar que el Sur contrafáctico habría estado mejor que el real en *todas* las asignaciones de la canasta familiar de los primeros con respecto a los miembros de los segundos. De todas maneras, pienso que en los casos interesantes los más altos ingresos de un estado serían más altos que los más bajos ingresos de los otros, por lo que, en la práctica, es imposible hacer comparaciones sobre esta base que no sean ambiguas. Por otra parte acabo de argumentar que no es posible encontrar una asignación privilegiada que preserve la identidad personal; en particular, las asignaciones que preservan los rangos, en el caso general no preservan la identidad personal de algún subgrupo independientemente definido.

Aun si evitáramos hacer comparaciones de subgrupos definidos funcionalmente podríamos tratar de tomar los problemas de las entidades agregadas, tal como el PBN *per capita* o las tasas de crecimiento. Sin embargo, en tal caso chocaríamos con el obstáculo del racismo. La relación causal entre la esclavitud y el racismo, el carácter persistente del racismo una vez creado (aun en ausencia de la causa original) y los efectos negativos del racismo sobre el desempeño de la economía, requiere un punto de ramificación tan atrás en el tiempo que la reconstrucción contrafáctica cruza el límite entre la ciencia y la ciencia ficción. El resultado de mi discusión del primer problema contrafáctico, es decir, el impacto de la

esclavitud sobre el crecimiento económico del Sur, es pues completamente negativo. El gran problema —el impacto de la esclavitud sobre el PBN o la tasa de crecimiento de la economía— no puede resolverse por razones prácticas; los problemas más pequeños —las consecuencias de la esclavitud para los subgrupos definidos en términos del sistema— no se pueden resolver por razones conceptuales. Invocar datos comparativos para probar que el Viejo Sur no podría haber tenido una tasa de crecimiento más alta sin la esclavitud, como hacen Fogel y Engerman,[68] no es muy relevante. En particular, hay una compensación entre los bajos niveles de producto bruto nacional *per capita* y las tasas de crecimiento relativamente altas, por lo menos en ciertos estadios del desarrollo económico, que podría igualmente invocarse para probar que sin la esclavitud, el Sur de 1860 habría tenido un ingreso *per capita* más alto.

Me dedicaré ahora al segundo problema contrafáctico, concerniente a qué hubiera ocurrido con la economía del Sur si no hubiera habido una Guerra Civil. Esta pregunta se encuentra inextricablemente ligada al análisis de la racionalidad, la ganancia y la viabilidad de la esclavitud. En los años recientes, esta discusión se ha hecho cada vez más compleja y sólo puedo ofrecer un esbozo simplificado de los argumentos. Enfocaré el problema a través de una diferencia entre las distintas tasas de ganancia (o de interés) que son conceptualmente importantes para el análisis. Estas son:

> r_1 = la tasa de interés prevaleciente que los inversores en 1860 obtuvieron fuera de la esclavitud, por ejemplo, en los ferrocarriles y las manufacturas;
> r_2 = la tasa de interés que habría prevalecido si la esclavitud hubiera desaparecido;
> r_3 = la tasa interna de retorno de la cría de esclavos, calculada por el valor de las ganancias brutas futuras esperadas al momento del nacimiento de un esclavo y el valor de los costos de mantenimiento estimados para su vida (asumiendo, para simplificar la cuestión, neutralidad de riesgo);
> r_4 = la conocida tasa de ganancias de la producción de los esclavos, definida como la proporción que ecualiza los precios corrientes de los esclavos y el valor presente de las ganancias netas esperadas para el futuro, estando tanto los precios como las ganancias definidos en relación con un esclavo de una edad dada (todavía se supone neutralidad de riesgo).

Para determinar las dos últimas tasas, necesitamos algunas ideas sobre las expectativas de futuras ganancias que tienen los dueños de esclavos. No sirve simplemente extrapolar las ganancias actuales hacia el futuro, porque entonces sólo obtenemos el caso especial de que los dueños de esclavos no anticiparon cambios en las ganancias. (Las ganancias actuales son relevantes sólo para la pregunta de si los precios *pasados*, basados en expectativas pasadas, fueron justificados o no.) Más bien debemos seguir a Fogel y Engerman, que utilizan los datos disponibles

para medir la "sanguinidad de los dueños de esclavos" mirando los movimientos relativos de las tasas de pago en relación a los esclavos y los precios de los esclavos. Un aumento en la tasa de pago en relación a los precios indicaría que los plantadores tenían un punto de vista pesimista con respecto a las ganancias futuras.[69]

Si los dueños de esclavos fueran "racionales" en el sentido de Fogel y Engerman, es decir, se movieran solamente por el deseo de maximizar el retorno sobre su inversión, entonces deberíamos tener $r_4 = r_1$ en equilibrio. Alfred Conrad y John Meyer en su contribución al problema[70] exigían $r_4 = r_2$, pero la mayor parte de los autores posteriores argumentaron que para un análisis motivacional del plantador individual este enfoque de equilibrio general de la contrafactibilización de conjunto no es relevante. Aunque fuera el caso de que todos los plantadores hubieran estado en mejor situación abandonando todos la esclavitud, la discusión del capítulo 5 demostró que esto podría haber tenido un impacto motivacional sólo si (i) todos los plantadores reconocieran el hecho, reconocieran que era reconocido por todos, etc., y (ii) no se pudiera obtener ninguna ventaja individual de ser el único que no dejó la esclavitud. Asumido esto (además de asumir que no le daba ventajas al plantador individual el ser el único que abandonara la esclavitud, lo que está fundamentado por el hecho de la esclavitud misma), la situación aparecería a grandes rasgos como un Juego del Seguro. La dificultad, por supuesto, reside en suponer (i); es por cierto poco probable que las condiciones estructurales para la información requerida estuvieran presentes.

Si $r_4 < r_1$, esto podría ser un signo de la irracionalidad de los dueños de esclavos, en el sentido de que eran movidos por consideraciones distintas de las puramente comerciales. De acuerdo con la interpretación de Fogel y Engerman, estos motivos no comerciales son el consumo conspicuo, es decir, el orgullo y el prestigio ligado a la posesión de esclavos. En su análisis de esta noción parecen cometer dos errores conceptuales que discutiré brevemente. Primeramente, argumentan que:

> Demostrar que había una correlación positiva entre la posesión de esclavos y el prestigio no liquida la cuestión de la causalidad. ¿El precio de los esclavos era alto porque la propiedad de los esclavos traía prestigio, o la posesión de esclavos traía prestigio porque su precio era alto? Para distinguir entre estas alternativas uno necesita averiguar si la tasa de rentabilidad esperada de los esclavos estaba por encima o por debajo de las tasas alternativas.[71]

Como los autores encuentran que la rentabilidad sobre los esclavos era comparable a la rentabilidad sobre otras inversiones, optan por la segunda alternativa: el prestigio era el efecto más que la causa del alto costo de los esclavos. Este argumento parece descansar sobre una confusión entre el razonamiento del equilibrio general y el del equilibrio parcial. El prestigio atribuido a los bienes altamente productivos es seguramente un fenómeno de equilibrio general, en el sentido de que un dueño de esclavos

derivaría prestigio de sus esclavos no porque *sus* esclavos fueran especialmente productivos, sino porque los esclavos *en general* eran altamente productivos. Desde el punto de vista del plantador individual (equilibrio parcial) hubiera sido por lo tanto deseable aumentar su posesión de esclavos más allá del punto de maximización de ganancias, porque la caída en la eficiencia productiva de sus esclavos llevaría consigo sólo una pérdida infinitesimal de prestigio. Como todos los terratenientes presumiblemente actuarían de la misma manera, habría una pérdida general de eficiencia hasta el punto en que todo el prestigio hubiera desaparecido. En este punto el terrateniente individual preferiría reducir su control sobre los esclavos; cuando cada uno hiciera esto el prestigio de la posesión de esclavos volvería a subir, lo que induciría a que el ciclo comenzara una vez más. Estamos, en realidad, tratando con un juego sin solución (no cooperativo).

En segundo lugar, y ésta es una dificultad más fundamental de su enfoque, Fogel y Engerman toman como seguro que la única alternativa a la maximización de ganancias es el modelo de "consumo conspicuo"; de todas maneras, parecen tomar como una cuestión normal que las demás alternativas implicarían desviaciones de la maximización de ganancias. Creo que existen otras alternativas a la maximización de ganancias, que en el nivel de la conducta podrían dar los mismos resultados observados con la maximización de ganancias. (Aquí, como en otros casos, no está justificada la deducción de las motivaciones de maximización de ganancias a partir de la conducta de maximización de ganancias.)[72] Esta es la dialéctica hegeliana entre el amo y el esclavo, discutida en el capítulo 4. En este modelo, el poder del amo sobre el esclavo es inseparable de su consumo conspicuo de los objetos producidos por los esclavos (que no es la misma cosa que el consumo conspicuo *de* esclavos). Para que el dueño derive prestigio social de sus esclavos no es suficiente con la mera posesión; debe también tener un nivel correlativamente alto de consumo. Para simplificar, podemos suponer que los esclavos y las mercancías de consumo ordinario deben mantenerse en proporciones fijas, para que con un número dado de esclavos un aumento en el consumo de mercancías más allá de la magnitud proporcional no dé ninguna satisfacción extra y viceversa. Las proporciones pueden ser constantes, para que las esquinas de las curvas de indiferencia se encuentren sobre una línea recta, pero en general tenemos tan sólo que suponer que más esclavos demandan más bienes de consumo. Pero si el consumo de los plantadores es una función cada vez mayor de la cantidad de esclavos utilizados en las plantaciones, esto significa que *el consumo de los terratenientes puede ser clasificado analíticamente como costos de la producción.* Eugene Genovese ha sido muy criticado por decir que los gastos de los plantadores en vacaciones, educación de sus hijos, etc. debían verse como costos de producción (a ser deducidos de las ganancias)[73], pero admito que su procedimiento puede justificarse de acuerdo con las líneas hegelianas ya señaladas. En una economía esclavista, Hegel puede ser leído diciendo que la escala de consumo está directamente ligada a la escala de producción, de tal modo que ambas están determinadas simultáneamente

antes que sucesivamente, mientras que en el análisis estándar del empresario capitalista se supone que el consumo depende del ingreso neto, los precios de las mercancías y las preferencias, siendo indiferente el volumen bruto de la producción.[74] De acuerdo con el modelo hegeliano, $r_4 = r_1$ no es condición suficiente para la racionalidad de los plantadores, ya que es bastante posible que la función que ligue el consumo a la cantidad de esclavos los hiciera querer mantener exactamente la cantidad de esclavos que maximizará las ganancias.

Con estos comentarios no quiero decir que he refutado lo que Fogel y Engerman afirman sobre la racionalidad de los dueños de esclavos. Sus conclusiones pueden ser correctas, aun a pesar de los agujeros en sus argumentos. Sólo digo que he demostrado (i) que la racionalidad es incompatible con el prestigio como un efecto de los altos precios de los esclavos, y no sólo con el prestigio como la causa de los altos precios de los esclavos, y (ii) que del hecho observado de que $r_4 = r_1$ no se puede concluir automáticamente que los dueños de esclavos fueran racionales, aun si esta conclusión fuera válida y el modelo de consumo conspicuo fuera la única alternativa a la maximización de ganancias. Dado el factor $r_4 = r_1$, podemos concluir sólo que la esclavitud era beneficiosa. Si Fogel y Engerman cometen el error de concluir la racionalidad de la ganancia, ellos deben recibir el crédito, a pesar de todo, de haber evitado concluir la viabilidad a partir de la ganancia. Ahora me dedicaré a ese problema.

La rentabilidad de la esclavitud significa sólo que no se observó ni sobrecapitalización ni baja capitalización de los retornos netos de los esclavos, pero con esta definición cualquier empresa moribunda va a parecer provechosa.[75] Dado el flujo esperado de retornos netos, no importa cuán pequeño (pero positivo), y la tasa prevaleciente de interés r_1, automáticamente el mercado fijará el valor real de los bienes productivos para demostrar la misma ganancia que los demás bienes. (Aquí supongo ganancia *y* racionalidad, para demostrar que juntas no implican viabilidad.) Por lo tanto, el problema de la viabilidad de la esclavitud debe definirse en términos del costo de la producción de esclavos más que en términos de los precios de los esclavos. Para un esclavo de 20 años, el valor presente de las ganancias netas futuras (descontado r_1) puede ser positivo aun si el valor al nacimiento de sus futuras ganancias netas (descontadas de r_1) es negativo. Esto es así porque hasta la edad de los diez años los gastos en un esclavo superan el ingreso producido por él, de tal modo que el flujo de retornos netos tendrá algunos elementos negativos que pueden contrapesar los elementos positivos, especialmente si los primeros, correspondientes a los primeros años y reducidos por un coeficiente de mortalidad menor, cuentan más que los segundos en el valor actual descontado. Decir que el valor al nacimiento de un esclavo es negativo cuando lo descontamos de la tasa de interés prevaleciente ($r_3 < r_1$) significa que la esclavitud no es viable. Fogel y Engerman hallaron que esto había ocurrido en Jamaica, donde la esclavitud puede decirse que no fue viable en el sentido de requerir un flujo

continuo de esclavos adultos desde Africa, mientras que en los Estados Unidos encontraron que $r_3 > r_1$.

En la medida en que la esclavitud era viable en este sentido *ex ante*, podemos también justificar la presunción de que no tuvo lugar la Guerra Civil y luego preguntarnos si podría haber sido viable en el sentido *ex post*. La cuestión aquí es la siguiente. Algunas interpretaciones tradicionales de la esclavitud han argumentado que la Guerra Civil fue causada por las contradicciones internas de la esclavitud; que los dueños de esclavos tenían la elección subjetiva entre una guerra expansionista y una declinación constante en las perspectivas. Si la esclavitud hubiera resultado ser no viable en el sentido *ex ante* (subjetivo), esta interpretación no habría sido considerada fuera de toda posibilidad. Y si aceptáramos que la causa principal de la guerra fueron las contradicciones económicas en el Sur antes que el sentimiento antiabolicionista en el Norte, entonces la guerra no podría ser tratada como un hecho exógeno a los estados esclavistas y no hubiera estado justificado asumir su ausencia (con el propósito de predecir el curso futuro del desarrollo económico del Sur). Por otra parte, con dueños de esclavos optimistas, parece legítimo asumir esto, y entonces podemos preguntarnos si su optimismo estaba en realidad justificado.

Como dije anteriormente, espero que el uso de estos largos ejemplos esté justificado por el interés que presentan por derecho propio, además de su importancia para el análisis conceptual de los contrafácticos históricos. Tal vez puede ser útil si en este punto resumo brevemente algunas de las ideas principales. He argumentado a favor de una interpretación específica de los contrafácticos históricos y en contra de la interpretación general de todos los contrafácticos propuesta por David Lewis. No he tratado de ofrecer una teoría general alternativa. Mis objeciones a la teoría de Lewis han sido de dos tipos. En primer lugar, en forma bastante general y sin referencia a los contrafácticos históricos, he argumentado que explicar la causalidad por medio de los contrafácticos y los contrafácticos a través de la similitud es un proceso circular, ya que la importancia causal es un elemento en nuestras nociones intuitivas sobre la similitud. En segundo lugar, he argumentado que aun si omitiéramos esta objeción y tratáramos de buscarle un sentido intuitivo a la noción de similitud, la explicación de Lewis es bastante engañosa cuando se la aplica al problema específico de los contrafácticos históricos, y que como se supone que es una teoría general para todo tipo de contrafácticos, esta sería una crítica de bastante importancia. El problema principal es que, de acuerdo con la teoría de Lewis, el mundo posible más cercano que satisface cierta condición puede muy bien resultar ser un mundo que nunca podría haberse desprendido del pasado real; por cierto, si nos restringimos a los mundos en los que las leyes de la naturaleza, aun si fueran diferentes de las nuestras, sean deterministas en forma, no podría haber ningún tipo de desprendimiento o ramificación. Sin embargo, cuando el historiador hace una afirmación contrafáctica, admito que se la pretende, y debe ser analizada como una afirmación sobre

lo que podría haber ocurrido (según lo que creemos) en el pasado *real*. Esto implica la necesidad de una teoría genética de los contrafácticos y de una teoría de la afirmabilidad más que de la verdad. Pienso que la teoría se ve fortalecida por su congruencia con la explicación genética de las propiedades esenciales adelantada por Saul Kripke y Michael Dummett. En lugar de buscar contrapartes de los mundos reales o de los individuos en ellos, debemos interpretar a los contrafácticos, sobre los actores individuales o colectivos, como afirmaciones acerca de *ellos*.

Notas

1. Lewis (1973a).

2. Para una de tales versiones, véase Mackie (1965); para objeciones véase Sosa (1975).

3. Unas pocas palabras sobre el método comparativo pueden ser útiles en este momento. En algunos casos privilegiados, la importancia relativa de las causas puede determinarse directamente sobre bases comparativas. Cuando argumentamos que la causa A fue más importante que la causa B para el surgimiento del efecto C en el país I, no siempre tenemos que comparar un país contrafáctico I sin A, un país contrafáctico I sin B y el país real I; con una gran suerte puede sernos posible solucionar la cuestión comparando el país real I con el país real II exactamente similar (en los aspectos definidos por la teoría subyacente) al país I excepto por la ausencia de A y con el país real III similar al país I excepto por la ausencia de B. En el caso general, sin embargo, las comparaciones juegan un rol más indirecto (aunque absolutamente crucial), como la base empírica para las generalizaciones que se necesitan para evaluar los contrafácticos. La evidencia comparativa puede utilizarse, por ejemplo, para establecer la generalización de que en las sociedades tradicionales la curva de la oferta de trabajo es negativa; en el análisis de una sociedad tradicional dada, para la cual no tenemos evidencia directa sobre este punto, esta generalización puede entonces utilizarse para determinar el efecto que hubiera tenido sobre la producción, de haber ocurrido, un aumento de salarios; y la respuesta a este problema puede a su vez decirnos algo sobre la importancia relativa del nivel de salarios comparado con las demás causas que influyen sobre el nivel del producto bruto nacional. Puede resultar útil poner el acento, para una referencia posterior, en que las proposiciones contrafácticas utilizan la ley y no la confirman. Si, sobre la base de ciertas generalizaciones, se afirma una proposición contrafáctica, esto no constituye una mayor prueba de la generalización. Pensar que sí lo es sería exponerse a la devastadora crítica de Adler emitida por Popper (1963, pág.35): "Una vez, en 1919, le informé un caso que no me pareció especialmente adleriano, pero que él no tuvo dificultad en analizar en términos de su teoría de los sentimientos de inferioridad, aunque nunca había visto al niño. Un poco sorprendido, le pregunté cómo podía estar tan seguro. «Debido a mis mil experiencias», replicó, ante lo cual no pude evitar decir: «Y con este nuevo caso, supongo que sus experiencias son mil una.»"

4. Leibniz (1948, págs. 285-286).

5. Rescher (1967, p.17); véase también Mates (1968).

6. Lewis (1968), Mondadori (1973).

7. Mill (1967, pág. 575).

8. Mill (1967, pág. 256).

9. Mill (1967, pág. 575).

10. Mill (1967,págs. 80-81).

11. Weber (1968, pág. 276).

12. Para afirmaciones similares, véase Hegel (1970, vol.16, pág. 290) y Merleau-Ponty (1963, pág. 45).

13. Lewis (1973b, págs. 66 y sigs.).

14. Aquí sigo a Ashby (1971, pág. 130).

15.Véase Berge (1963) para una exposición.

16.Un ambicioso intento (el único por lo que sé) de formular una teoría del cambio histórico en estos términos es Rader (1971).

17.Véase aquí Davidson (1973), como fue parafraseado en el apéndice 2 del capítulo 5; también Elster (1976a).

18.Este, supongo, es el punto que establece McClelland (1975, págs. 161 y sigs.) cuando enfatiza el rol esquivo pero importante de los sustitutos que podrían haber surgido en lugar del factor mentalmente quitado de la presuposición contrafáctica.

19. Goldman (1972, pág. 223).

20.Stalnaker (1968), Lewis (1973b).

21.En realidad esta simplificación de la teoría de Lewis nos da la formulación original de Stalnaker. Lewis quiere mantener abierta la posibilidad de que pueda haber un mundo no A que fuera *el* más cercano, ya sea porque hay mundos A arbitrariamente cercanos al nuestro o porque puede haber varios mundos A igualmente y mínimamente distantes del nuestro. La segunda idea es la más importante; sin ella se desploma la diferencia entre los contrafácticos-sería y los contrafácticos-podría.

22. Nagel (1963, pág. 214).

23.Mondadori (1973); la cita no es exacta.

24. Además Lewis (1973a) *cita el problema de los efectos*: asumiendo que *c* sea causa suficiente de *e* (dadas las leyes y algunas de las circunstancias reales) debemos concluir que si *e* no hubiera ocurrido, entonces *c* no habría ocurrido, por lo que *e* resulta ser causa de *c*. Este problema se resuelve por el mismo método que el problema de los epifenómenos.

25. Lewis (1973a, pág. 566)

26. Véanse los comentarios al fin de la nota 3.

27. Climo y Howells (1976).

28. Climo y Howells (1976, págs. 11-12).

29. Lewis (1973a, pág. 567).

30. Climo y Howells (1976, pág. 14).

31. La manera obvia de definirla relación de accesibilidad entre dos mundos actualizados es decir que s'_t es posible en relación a s''_t si hay algunos estados s_{t1} en el pasado real a partir del cual podrían haber emergido s'_t y s''_t, al final de las trayectorias permitidas. No es obvio, sin embargo, qué medida de la distancia entre s'_t y s''_t sería la apropiada. Dado que s'_t y s''_t están a distancias d' y d'', respectivamente, del estado real s_{t1}, entonces $(d'+d'')$, / $d''-d''$/ y max (d',d'') son todos candidatos posibles para la medida de la distancia entre s'_t y s''_t.

32. Simensen (1977).

33. Manning (1974).

34. La siguiente explicación se toma en gran medida de Elster (1975b).

35. El término "optimismo" fue acuñado en 1737 para denotar la doctrina leibniziana de que el mundo real es el mejor de todos los mundos posibles; por lo tanto, se puede aplicar este término al aspecto contrafáctico de la controversia. El término "pesimismo" fue utilizado por Schopenhauer en 1819 para denotar su propia doctrina de que el mundo real es el peor de todos los mundos posibles, pero una aceptación más común es la idea de que el mal en el mundo supera al bien, que es una comparación intramundo similar a la denotada aquí por el mismo término.

36. Hartwell y Engerman (1975).

37. Rawls (1971, pág. 98).

38. Barry (1973, pág. 50).

39. Compárese Okun (1975).

40. Hobsbawm (1963, pág. 112).

41. Así la siguiente afirmación está abierta a dudas: "El trabajador del campo sufrió más agudamente bajo la presión de la crisis industrial, aunque ganaba sustancialmente lo *mismo* cuando la actividad de los negocios aumentaba" (Taylor, 1960, pág. 29, las itálicas son mías).

42. Hobsbawm (1957, pág.46), criticado por Hartwell (1961, pág.413, n.2).

43. "Es necesario tomar en cuenta no sólo a los años de prosperidad y depresión sino también tal vez a la nueva inseguridad que el cambiante carácter del ciclo de negocios trajo consigo" (Taylor, 1960, pág.29). Contra esto podemos notar la preferencia de riesgo de los muy pobres.

44. Para afirmaciones recientes en contrario véase Pollak (1976b) y Haavelmo (1970); también Leibenstein (1976, capítulo 4) .

45. Por favor, obsérvese que no es lo mismo que la dependencia de mis preferencias actuales del consumo pasado, como lo expliqué en la sección sobre el cambio endógeno de los gustos, en el capítulo 2.

46. Hayek (1954, pág.16).

47. Hartwell (1961, pág.412); Habakkuk (1971, pág34).

48. Henderson y Chaloner (1958, pág. xiv).

49. Hobsbawm (1963, págs. 127-128).

50. Marx (1933, pág.120); para Rosa Luxemburgo, véase la discusión en el capítulo 5 y la referencia en la nota 133 de ese capítulo.

51. Hayek (1954, pág. 28).

52. Loschky (1973).

53. Walton (1971).

54. Este es el argumento de Thomas (1965, pág. 637, n.53).

55. Thomas (1965) criticado por Walton (1971, pág.537).

56. Fogel (1964).

57. Schmookler (1966).

58. Podemos observar aquí que la distinción entre los aspectos funcionales y materiales de un invento permiten una simple prueba de la tesis de que es imposible predecir el desarrollo tecnológico. En un tiempo t_1 es lógicamente imposible predecir los detalles materiales de los inventos que serán hechos en t_2, ya que una predicción exitosa sería equivalente a hacer el invento. Podemos asumir, por lo menos en favor del argumento, que es posible predecir en t_1 los aspectos funcionales de los inventos que se harán en t_2, pero no podemos predecir en t_1 los aspectos funcionales o materiales de los inventos que se harán en t_3, ya que algunos de los mismos surgirán directamente de los (impredecibles) aspectos materiales de los inventos de t_2 y, por lo tanto, tan imposibles de conocer en t_1 como lo son los segundos.

59. Merton (1973, Parte 4); Kroeber (1952).

60. Rosenberg (1974).

61. David (1969).

62. Fogel y Engerman (1974).

63. Stampp (1976).

64. Fogel y Engerman (1975, vol.II, pág.76).

65. El término es tomado de Georgescu-Roegen (1971, pág.126), quien pregunta "¿Puede crearse el hombre socialista de tal modo que no muestre ningún rastro histéresis de su pasado burgués o campesino?" Véase también Elster (1976a).

66. Russel (1966).

67. Fogel (1967, pág.287).

68. Fogel y Engerman (1974, vol.I, págs 247 y sigs). Los autores argumentan que el Sur había tenido tanto un alto ingreso *per capita* como altas tasas de crecimiento comparado con la mayoría de los demás países, pero la comparación más pertinente parecería ser con el Norte, donde el ingreso *per capita* era el 40% más alto que en el Sur, aun si la tasa de crecimiento era sólo el 82% de la del Sur.

69. Para estos problemas, con referencias completas a la literatura, véase Fogel y Engerman (1971) y Fogel y Engerman (1974, vol.II, págs.54 y siguientes).

70. Conrad y Meyer (1958).

71. Fogel y Engerman (1971, vol. I, pág. 71).

72. Véase David y Temin (1976, pág.40) para una observación similar.

73. Genovese (1965m pág. 278), criticado por A.H. Conrad (1967, pág.520) y Fogel y Engerman (1971, pág. 321).

74. Una excepción es Marx (1867, pág. 594): "Cuando un cierto estadio de desarrollo del capitalismo ha sido alcanzado, un grado convencional de prodigalidad, que es también una exhibición de riqueza, y consecuentemente una fuente de crédito, se hace una necesidad para el capitalista «desafortunado». El lujo entra dentro de los gastos de representación del capital." Para detalles del modelo Hegel-Genovese, véase Elster (1976c).

75. De aquí en adelante sigo a Yasuba (1961), así como también los posteriores desarrollos de su trabajo hechos por Fogel y Engerman, en el trabajo citado en la nota 69.

Referencias bibliográficas

Acton, H.B. (1967). "Dialectical materialism", *The Encyclopedia of Philosophy*, Nueva York: Macmillan.

Ainslie, G. (1975). "Specious reward", *Psychological Bulletin*, 82, 463-496.

Allen, R.G.D. (1966). *Mathematical Economics*, 2ª ed., Londres: Macmillan. [Versión castellana: *Economía matemática*. Madrid, Aguilar.]

Althusser, L.(1965). *Pour Marx*, París: Maspero.

Arrow, K. (1971a). *Essays in the Theory of Risk-Bearing*, Amsterdam: North-Holland.

Arrow, K. (1971b). "Political and economic evaluation of social effects and externalities", en Intriligator, M.D. (comp.), *Frontiers of Quantitative Economics*, Amsterdam: North-Holland.

Arrow, K. (1974). *The Limits of Organization*, Nueva York: Norton .

Arrow, K. y Scitovsky, T. (comps.) (1969). *Readings in Welfare Economics*, Londres: Allen y Unwin.

Ashby, W.R. (1960). *Design for a Brain*, Londres: Chapman and Hall.

Ashby, W.R. (1971). *Introduction to Cybernetics*, Londres: Chapman and Hall.

Avineri, S. (1969). *The Social and Political Thought of Karl Marx*, Cambridge: Cambridge University Press. [Versión castellana: *El pensamiento social y político de Carlos Marx*. Madrid, Centro de Estudios Constitucionales, 1983.]

Avineri, S. (1971). "Consciousness and history: *List der Vernuft* in Hegel y Marx", en W.E Steinhaus (comp.), *New Studies in Hegel's Philosophy*, Nueva York: Holt, Rinehart, and Winston.

Balacz, E. (1968). *La Bureaucratie Celeste*, París: Gallimard.

Bachrach, P. y Baratz, M. (1963), "Decisions and nondecisions", *American Political Science Review*, 57, 632-642.

Balibar, E. (1966). "Les concepts fondamentaux du matéralisme historique", en L.Althusser y otros, *Lire le Capital*, vol.II, París: Maspero. [Versión castellana: *Cinco ensayos de materialismo histórico*. Barcelona, Editorial Laia, 1976.]

Bar-Hillel, Y., Perles,M., y Shamir, F. (1964). "On formal properties of simple phrase structure grammars", en Y.Bar-Hillel, *Language and Information*, Conferencia,Mass.: Addison-Wesley.

Barnes, J. (1971). *Three Styles in the Study of Kinship*, Londres: Tavistock.

Barry, B. (1973). *The Liberal Theory of Justice*, Oxford: Oxford University Press.

Bateson, G. (1956). "Towards a theory of schizophrenia", *Behavioral Science*, 1, 251-264.

Bateson, G. (1972). *Steps to an Ecology of Mind*, Nueva York: Ballantine Books.

Becker, G. (1962). "Irrational behavior in economic theory", *Journal of Political Economy*, 70, 1-13.

Benedict, B. (1972). "Social regulation of fertility", en G.A. Harrison y A.J. Boyce (comps.), *The Structure of Human Populations*, Oxford: Oxford University Press.

Berge, C. (1963). *Topological Spaces*, Edimburgo: Oliver and Boyd.

Bernstein, E. (1899). *Die Voraussetzungen des Sozialismus*, citado aquí de acuerdo con la edición de 1967, Hamburgo: Rowohlt. [Versión castellana: *Socialismo democrático*. Madrid, Tecnos, 1990.] [Del mismo autor: *Socialismo evolucionista*. Barcelona, Fontamara, 1974.]

Bjørgum, J. (1976). "*Fagopposisjonen av 1911*", *Tidsskrift for Arbeiderbevegelsens Historie* (Oslo), No. 1, 63-131.

Blau, P. y Duncan,O.D. (1967). *The American Occupational Structure*, Nueva York: Wiley.

Bliss, C.J. (1975). *Capital Theory and the Distribution of Income*, Amsterdam: North-Holland.

Bloch, E.(1954). *Das Prinzip Hoffnung*, Berlin: Aufbau.

Böhm-Bawerk, E. von (1898), *Karl Marx and the Close of his System*, Londres: Unwin.

Boudon, R. (1973). *Mathematical Structures of Social Mobility*, Amsterdam: Elsevier.

Boudon, R. (1974). *Education, Opportunity and Social Inequality*, Nueva York: Wiley.

Boudon, R. (1977). *Effects Pervers et Ordre Social*, París: Presses Universitaires de France.

Bourdieu, P., Chamboredon, J.-C. y Passeron, J.-C- (1968). *Le Métier de Sociologue*, París: Mouton. [Versión castellana: *El oficio de sociólogo*. Madrid, Siglo XXI, 2ª ed., 1989.]

Brams, S. (1976). *Paradoxes in Politics*, Nueva York: Free Press.
Bródy, A. (1974). *Proportions, Prices and Planning*, Amsterdam: North-Holland.
Bronfenbrenner, M. (1965). "Das Kapital for the modern man", *Science and Society*, 29, 419-438.
Brooke, M.(1972). "Problems in the decision-making process", en M.Brooke y H. Remmers (comps.), *The Multinational Company in Europe*, Londres: Longman.
Canarella, G. y Tomaske, J. (1975). "The optimal utilization of slaves", *Journal of Economic History*, 35.
Capek, M. (1972). "The fiction of instants", en J.T. Fraser y cols. (comps.), *The Study of Time*, Berlin: Sringer.
Chamley, P.(1963). *Economie Politique et Philosophie chez Steuart et Hegel*, París: Dalloz.
Chayanov, A.V. (1966). *The Theory of Peasant Economy*, Homewood, Ill.: Irwin.
Chomsky, N. (1957). *Syntactic Structures*, La Haya: Mouton.
Chomsky, N. (1959). "On certain formal properties of grammars", *Information and Control*, 2, 137-167.
Chomsky, N. (1964). "Current issues in Linguistic theory", en J.A. Fodor y J.J. Katz (comps.), *The Structure of Language*, Englewood Cliffs, NJ: Prentice-Hall.
Chomsky, N. y Miller, G.A. (1963). "Finitary models of language users", en R.D. Luce, R.R. Bush y E. Galanter (comps.) , *Handbook of Mathematical Psychology*. vol. II, Nueva York: Wiley.
Church, A. (1956). *Introduction to Mathematical Logic*, Princeton: Princeton University Press.
Clark, J.M. (1899). *The Distribution of Wealth*, Nueva York: Macmillan.
Climo, T.A. y Howells, P.G.A. (1976). "Possible worlds in historical explanation", *History and Theory*,15.
Cobb, W.E.(1973). "Theft and the two hypotheses", en S.Rottenberg (comp.), *The Economics of Crime and Punishment*, Washington: American Enterprise Institute for Public Policy Research.
Coddington, A. (1968). *Theories of the Bargaining Process*, Londres: Allen and Unwin.
Cole, J. y Cole, S. (1973). *Social Stratification in Science*, Chicago: Chicago University Press.
Coleman, J. (1973a). *The Mathematics of Collective Action*, Londres: Heinemann.
Coleman, J. (1973b). "Theoretical bases for parameters of stochastic processes", en P.Halmos (comp.), *Stochastic Processes in Sociology*, Monografías de *Sociological Review*, No.*19*.
Conrad, A.J. (1967). Contribución al "Symposium on slavery as an obstacle to economic growth", *Journal of Economic History*, 27, 518-560.
Conrad, A.J. y Meyer, J. (1958). "The economics of slavery in the ante-bellum south", *Journal of Political Economy*, 66, 95-122.
Cripps, T.F. y Tarling, R.J. (1974). "An analysis of the duration of male unemployment in Great Britain 1932-63". *Economic Journal, 84*, 289-316.
David, P. (1969). "Transport innovation and economic growth" *Economic History Review*, 22, 506-525.
David, P. (1975). *Technical Choice, Innovation and Economic Growth*, Cambridge: Cambridge University Press.
David P. y Temin, P. (1976). "Capitalist masters, bourgeois slaves", en P.David, y cols., *Reckoning with Slavery*, Nueva York: Oxford University Press.
Davidson, D. (1963). "Actions, reasons and causes", *Journal of Philosophy*, 60, 685-700.
Davidson, D. (1969). "How is weakness of the will possible?", en J.Feinberg (comp.), *Moral Concepts*, Oxford: Oxford University Press.
Davidson, D. (1970). Mental events, en L. Foster y J. W. Swanson (comps.), *Experience and Theory*, Cambridge, Mass.: University of Massachusetts Press.
Davidson, D. (1973). "The material mind", en P. Suppes y cols. (comps.), *Logic, Methodology and Philosophy of Science IV*, Amsterdam: North-Holland.
Deutsch, K. (1963). *The Nerves of Government*, Glencoe, Ill.: Free Press.
Dorfman, R., Samuelson, P., y Solow, R. (1958). *Linear Programming and Economic Analysis*, Nueva York: McGraw- Hill . [Versión castellana: *Programación lineal y análisis económico*. Madrid, Aguilar, 1969.]

Dubarle, D. (1970). "Logique formalizante et logique Hégélienne", en J.d'Hondt (comp.), *Hegel et la Pensée Moderne*, París: Presses Universitaires de France.
Dummett, M. (1973). *Frege: Philosophy of Language*, Londres: Duckworth.
Duncan, O.D. (1966). "Methodological issues in the analysis of social mobility", en N.J. Smelser y S.M.Lipset (comps.), *Social Structure and Mobility in Economic Development*, Chicago: Aldine.
Easton, D. (1965). *A Systems Analysis of Political Life*, Nueva York: Wiley. [Versión castellana: *Esquema para el análisis político*. Buenos Aires, Amorrortu Editores, 1979.]
Eckstein, A. (1968). "Economic Fluctuations in Communist China", en Ping-ti Ho y Tang Tsou (comps.), *China in Crisis*, vol. I, Chicago: Chicago University Press.
Ehrenfest, P. y T. (1959). *Conceptual Foundations of Statistical Mechanics*, Itaca, NY: Cornell University Press .
El-Safty, A.E. (1976). "Adaptive behavior, demand and preferences", *Journal of Economic Theory*, 13, 298-318.
Elster, J. (1972). "Production et reproduction". *Essai sur Marx, Thèse d'Etat* no publicada, Université de París V.
Elster, J. (1975a). *Leibniz et la Formation de l'Esprit Capitaliste*, París: Aubier-Montaigne.
Elster, J. (1975b). "Optimism and pesimism in the discussion of the standard of living during the Industrial Revolution in Britain", informe al XIV Congreso Internacional de Ciencias Históricas, San Francisco.
Elster, J. (1976a). "A note on hysteresis in the social sciences", *Synthese*, 33, 371-391.
Elster, J (1976b). "Boudon, education and the theory of games", *Social Science Information*, 15, 733-740.
Elster, J. (1976c). "Some conceptual problems in political theory", en B.Barry (comp), *Power and Political Theory*, Londres: Wiley.
Elster, J. (1976d). "Ulysses and meta-rationality", trabajo presentado ante el ECPR *Workshop on Political theory*, Lovaina.
Elster, J. (1977). "Beyond gradient-climbing", trabajo presentado al Cuarto Congreso Internacional de la International Organization for the Study of Human Development.
Fararo, T.S., (1973). *Mathematical Sociology*, Nueva York: Wiley.
Farquhar, J.E. (1964). *Ergodic Theory in Statistical Mechanics*, Londres: Wiley.
Feller, W. (1968). *An Introduction to Probability Theory and its Aplications*, 3ª ed., vol. 1, Nueva York: Wiley.
Fellin, P. y Litwak, E. (1963). "Neighborhood cohesion under conditions of mobility", *American Sociological Review*, 28, 364-376.
Fellner, W. (1961). "Two propositions in the theory of induced innovations", *Economic Journal*, 71, 305-308.
Festinger, L. (1957). *A Theory of Cognitive Dissonance*, Stanford: Stanford University Press. [Versión castellana: *Teoría de la disonancia cognoscitiva*. Madrid, Centro de Estudios Constitucionales, 1975.]
Festinger, L. (1964). *Conflict, Decision and Dissonance*, Londres: Tavistock.
Finley, M.I. (1973a). *The Ancient Society*, Londres: Chatto and Windus.
Finley, M.I. (1973b). *Democracy: Ancient and Modern*, Londres: Chatto and Windus.
Fogel, R.W. (1964). *Railroads and American Economic Growth*, Baltimore: John Hopkins Press. [Versión castellana: *Los ferrocarriles y el crecimiento económico de los Estados Unidos*. Madrid, Tecnos, 1972.]
Fogel, R.W. (1967). "The specification problem in economic history", *Journal of Economic History*, 27, 283-308.
Fogel, R.W. y Engerman, S. (1971). "The economics of slavery", in R.Fogel y S. Engerman (comps.), *The Reinterpretation of American Economic History*, Nueva York: Harper and Row.
Fogel, R.W. y Engerman, S. (1974), *Time on the Cross*, Boston: Little, Brown. [Versión castellana: *Tiempo en la cruz, la economía esclavista en los Estados Unidos*. Madrid, Siglo XXI, 1981.]
Forrester, J. (1969). *Urban Dynamics*, Cambridge, Mass: MIT Press.
Freedman, J.L. y Sears, D.O. (1965). "Selective exposure", en L.Bercowirtz (comp.), *Advances*

in Experimental Psychology, vol.2, Nueva York: Academic Press.
Freud, S. (1968). "Die Verneinung" en S.Freud, *Gesammelte Werke*, vol XIV, Francfort: Fischer.
Gallagher, J. y Robinson, R. (1963). "The imperialism of free trade", *Economic History Review*, 6, 1-15.
Geach, P. (1972), *Logic Matters*, Berkeley y Los Angeles: University of California Press.
Genovese, E. (1965), *The Political Economy of Slavery*, Nueva York: Pantheon Books. [Versión castellana: *Economía política de la esclavitud.* Barcelona, Edicions 62, 1970.]
Genovese, E. (1974), *Roll, Jordan, Roll*, Nueva York: Pantheon Books.
Georgescu-Roegen, N. (1950), "The theory of choice and the constancy of economic laws", *Quarterly Journal of Economics*, 64, 125-138.
Georgescu-Roegen, N. (1968), "Utility", en *The International Encyclopedia of the Social Sciences*, Nueva York: Macmillan.
Georgescu-Roegen, N. (1971), *The Entropy Law and the Economic Process*, Cambridge, Mass.:Harvard University Press.
Germani, G. (1966), "Social and political consequences of mobility", en NJ. Smelser y S.M. Lipset (comps.), *Social Structure and Mobility in Economic Development*, Chicago: Aldine.
Gerschenkron, A. (1965), "Agrarian policies and industrialization: Russia 1861-1917", en H.J.Habakkuk y M. Postan (comps.), *Cambridge Economic History of Europe*, vol VI, Cambridge: Cambridge University Press.
Gerschenkron, A. (1966), *Economic Backwardness in Historical Perspective*, Cambridge, Mass.: Harvard University Press.
Girod, R. (1971), *Mobilité Sociale* , Ginebra: Droz.
Glantz, S.A. y Albers, N.V. (1974). "Department of Defense R and D in the university", *Science*, 186, 706-711.
Godelier, M. (1966). "Système, structure et contradiction dans Le Capital", *Les Temps Modernes*, 22, 828-864.
Goldman, A. (1972). "Toward a theory of social power", *Philosophical Studies*, 23, 221-268.
Gorman, W.M. (1967). "Tastes, habits and choices", *International Economic Review*, 8, 218-222.
Graaff, J.de V. (1971), *Theoretical Welfare Economics*, Cambridge: Cambridge University Press. [Versión castellana: *Teoría de la economía del bienestar.* Buenos Aires, Amorrortu Editores, 1967.]
Granovetter, M. (1973), "The strength of weak ties", *American Journal of Sociology*, 78, 1360-1380.
Guilbaud, G.Th. (1952). "Les théories de l'intéret général et le problème logique de l'agrégation, *Economie Appliquée*, 5, 501-584.
Günther, G. (1959). *Idee und Grundriss zu einer nicht-Aristotelischen Logik*, Hamburgo: Felix Meiner.
Haavelmo, T. (1944). "The probability approach in econometrics", *Econometrica (Supplement)*, 12, 1-118.
Haavelmo, T. (1970), "Some observations on welfare and economic growth", en W.A. Eltis, M. Scott y N.Wolfe (comps.), *Induction, Growth and Trade: Essays in Honour of Sir Roy Harrod*, Oxford: Oxford University Press.
Haavelmo, T.(1971). "Forventninger og deres Rolle i Økonomisk Teori, memorandum fra Socialøkonomisk Institutt ved Universitetet i Oslo".
Habakkuk, H.J. (1971). *Population Growth and Economic Development since 1750*, Leicester: Leicester University Press.
Hamblin, C.L. (1970). *Fallacies*, Londres: Methuen.
Hammond, P.J. y Mirrlees, J.A. (1973). "Agreeable plans", en J.A. Mirrlees y N.H Stern (comps.), *Models of Economic Growth*, Londres: Macmillan.
Harcourt, G.C. (1973), *Some Cambridge controversies in the Theory of Capital*, Cambridge:

Cambridge University Press. [Versión castellana: *Teoría del capital.* Vilassar de Mar, Oikos-Tau, 1975.]
Hartwell, R.M. (1961). "The raising standard of living in England 1800-1850", *Economic History Review,* 13, 397- 416.
Hartwell, R.M. y Engerman S. (1975). "Models of immiseration: The theoretical basis of pessimism", en A.J.Taylor (comp.), *The Standard of Living in Britain in the Industrial Revolution*, Londres: Methuen.
Hayek, F.A. (1954). "Introduction to F.A.Hayek y cols., *Capitalism and the Historians,* Chicago: Chicago University Press. [Versión castellana: *El capitalismo y los historiadores,* Madrid, Unión Editorial, 1973.]
Hayek, F.A. (1967). "The results of human action but not of human design", en F.A. Hayek, *Studies in Philosophy, Politics and Economics*, Londres: Routledge.
Heal, G.M. (1973). *The Theory of Economic Planning*, Amsterdam: North-Holland, [Versión castellana: *Teoría de la planificación económica,* Barcelona, Bosch Casa Editorial, 1977.]
Hegel, G.W.F. (1961). *The Science of Logic*, Londres: Allen and Unwin. [Versión castellana: *Lógica.* Barcelona, Orbis, 1984.]
Hegel, G.W.F. (1970). *Werke in 20 Bänden*, Francfort: Suhrkamp.
Henderson, W.O. y Chaloner, W.H. (1958). "Introduction to F.Engels", en *The Conditions of the Working Class in England*, Oxford: Oxford University Press.
Hintikka, J. (1961). *Knowledge and Belief,* Ithaca, NY: Cornell University Press. [Versión castellana: *Saber y creer. Una introducción a la lógica de las dos nociones*, Madrid, Tecnos, 1979.]
Hintikka, J. (1967). "Cogito, ergo sum: Inference or Performance?", en W.Doney (comp.), *Descartes*, Londres: Macmillan. [Del mismo autor: *Ensayos sobre explicación y comprensión.* Madrid, Alianza Editorial, 1980.]
Hintikka, J. (1972). "Leibniz on Plenitude, relations and the 'Reign of law'", en H.G.Frankfurt (comp.), *Leibniz*, Nueva York: Anchor Books.
Hintikka, J. (1973). *Logic, Language-Games and Information*, Oxford: Oxford University Press. [Versión castellana:*Lógica, juegos de lenguaje e información,* Madrid, Tecnos, 1976.]
Hirschman, A.O. (1958). *The Strategy of Economic Development*, New Haven: Yale University Press.
Hirschman, A.O. (1970). *Exit, Voice and Loyalty,* Cambridge, Mass.: Harvard University Press.
Hirschman, A.O. (1973). "Changing tolerance for inequality", *Quarterly Journal of Economics,* 87, 544-566.
Hobsbawm, E. (1954). "The general crisis of the 17th- century", *Past and Present,* No. 5, Part I, 33-53; No.6, Part II, 44-65.
Hobsbawm, E. (1957). "The British standard of living 1790- 1850", *Economic History Review,* 10, 46-61.
Hobsbawm, E. (1963). "The standard of living during the Industrial Revolution", *Economic History Review*, 16, 120-134.
Hockett, C.F. (1955). *A Manual of Phonology*, Baltimore: Memoria 11 del *International Journal of American Linguistics.*
Hoel, M. (1974). "Kryssløp, Arbeidsverdier og Priser", *Statsøkonomisk Tidsskrift,* 88, 57-88.
d'Hont, J. (1970). "Téléologie et praxis dans la logique de Hegel", en J.d'Hont (comp.), *Hegel et la Pensée Moderne*, París: Presses Universitaires de France.
Howard, N. (1971). *Paradoxes of Rationality*, Cambridge, Mass.: MIT Press.
Hughes, G.E, y Cresswell, M.J. (1972). *An Introduction to Modal Logic*, Londres: Methuen. [Versión castellana: *Introducción a la lógica modal.* Madrid, Tecnos., 1973.]
Hull, D. (1965). "The effect of essentialism on taxonomy", *British Journal of Philosophy of Science*, 15, 314-326.
Hume, D. (1748). *An Enquiry into Human Understanding.* [Versión castellana: *Investigación sobre el conocimiento humano.* Madrid, Alianza Editorial, 8ª ed., 1992.]
Husserl, E. (1913). *Logische Untersuchungen*, vol I: Prolegomena, 2a. edición, Tübingen: Max Niemeyer. [Versión castellana: *Investigaciones lógicas.* Madrid, Alianza Editorial, 2ªed., 1985.]
Husserl, E. (1966). *Husserliana*, vol.X, Haag: Martinus Nijhoff.

Husserl, E. (1967). *Husserliana*, vol.XI, Haag: Martinus Nijhoff.
Hyppolite, J. (1946). *Genèse et Structure de la Phénoménologie de l'Esprit de Hegel*, París: Aubier- Montainge. [Versión castellana: *Génesis y estructura de la Fenomenología del espíritu de Hegel*. Barcelona, Edicions 62, 2ª ed., 1991.]
Hyppolite, J. (1966). "Commentaire parlé sur la 'Verneinung' de Freud", en J. Lacan, *Ecrits*, París: Editions du Seuil.
Ishiguro, H. (1972a), "Leibniz's theory of ideality of relations", en H.G. Frankfurt (comp.), *Leibniz*, Nueva York: Anchor Books.
Ishiguro, H. (1972b). *Leibniz's Philosophy of Logic and Language*, Londres: Duckworth.
Isnard, C.A. y Zeeman, E.C. (1974). "Some models from catastrophe theory in the social sciences" en L.Collins (comp.), *Use of Models in the Social Sciences*, Londres: Tavistock.
Jakobson, R. y Halle M., (1963). "Phonologie et phonétique" en R. Jakobson, *Essais de Linguistique Générale*, París: Editions de Minuit.
Johansson, I. (1936). "Der Minimalkalkül", *Compositio Mathematica*, 4, 119-136.
Kalecki, M. (1943). "Political aspects of full employment", *Political Quarterly*, 14, 322-331.
Kanger, S. y Kanger, H. (1966). "Rights and Parliamentarism", *Theoria*, 32, 85-115.
Kaplan, M. (1967). "Systems theory" en J.C. Charlesworth (comp.), *Contemporary Political Analysis*, Nueva York: Free Press.
Kemeny, J.G., Snell, J.L., y Thompson, G.L. (1966). *Introduction to Finite Mathematics*, 2a.edición, Englewood Cliffs, NJ: Prentice-Hall.
Kenny, A. (1963). *Action, Emotion and Will*, Londres: Routledge.
Kenny, A. (1965-66). "Happiness", *Proceedings of the Aristotelian Society*, 66 (N.S.), 93-102.
Kenny, A. (1970). "Intention and purpose in law", en R.S.Sumers (comp.), *Essays in Legal Philosophy*, Oxford: Blackwell.
Kerr, C. y Siegel, A. (1964). "The inter-industry propensity to strike", en C.Kerr, *Labor and Management in Industrial Society*, Nueva York: Anchor Books.
Keynes, J.M. (1936). *The General Theory of Employment, Interest and Money*, Londres: Macmillan. [Versión castellana: *Teoría general de la ocupación, el interés y el dinero*. Madrid, Fondo de Cultura Económica, 12ª ed., 1980.]
Keynes, J.M. (1971). *A Treatise on Money*, vol.V de *The Collected Works of John Maynard Keynes*, Londres: Macmillan.
Kindleberger, C. (1975). "The rise of free trade in Western Europe 1820-75", *Journal of Economic History*, 35, 20-55.
Körner, S. (1976). *Experience and Conduct*, Cambridge: Cambridge University Press.
Koyré, A. (1971). *Etudes de l'Histoire de la Pensée Philosophique*, París: Gallimard. [Versión castellana: *Estudios de historia del pensamiento científico*. Madrid, Siglo XXI, 3ª ed., 1990.]
Kripke, S. (1972). "Naming and necessity", en D.Davidson y G. Harman (comps.), *Semantics of Natural Language*, Dordrecht: Reidel, 253-355.
Kroeber, A. (1952). "The superorganic", en A.Kroeber, *The Nature of Culture*, Chicago: Chicago University Press.
Lacan, J. (1966). *Ecrits*, París: Editions du Seuil. [Versión castellana: *Escritos*. Buenos Aires, Siglo XXI, 13ª ed., 1985.]
Lancaster, K. (1973). "The dynamic inefficiency of capitalism", *Journal of Political Economy*, 81, 1092- 1109.
Laplanche, J. (1970). *Vie et Mort en Psychanalyse*, París: Flammarion. [Versión castellana: *Vida y muerte en psicoanálisis*. Buenos Aires, Amorrortu Editores, 1ª ed., 1973.]
Latsis, S. (1976). "A research programme in economics", en S.Latsis (comp.), *Methods and Appraisal in Economics*, Cambridge: Cambridge University Press.
Lave, C.A. y March J.G. (1975). *An Introduction to Models in the Social Sciences*, Nueva York: Harper and Row.
Lazarsfeld, P.A. y Menzel, H. (1969). "On the relation between individual and collective properties", en A. Etxioni (comp.), *A Sociological Reader on Complex Organizations*, Nueva York: Holt, Rinehart and Winston.
Lee, H.P.D. (1936). *Zeno of Elea*, Cambridge: Cambridge University Press.
Leibenstein, H. (1976). *Beyond Economic Man*, Cambridge, Mass.: Harvard University Press.

Leibniz, G.W. (1875-90). *Die Philosophische Schriften*, Gerhardt, Berlin: Weidmannsche Buchhandlung.
Leibniz, G.W. (1903). *Opuscules et Fragments Inédits*, Couturat, París: Presses Universitaires de France.
Leibniz, G.W. (1948). *Textes Inédits*, Grua, París: Presses Universitaires de France.
Leijonhufvud, A. (1976). "Schools, 'revolutions' and research programmes in economic theory", en S.Latsis (comp.) *Methods and Appraisal in Economics*, Cambridge: Cambridge University Press.
Lévi-Strauss, C. (1958). *Anthropologie Structurale*, París: Plon. [Versión castellana: *Antropología estructural.* Barcelona, Paidós Ibérica, 2ª ed., 1992.]
Lévi-Strauss, C. (1967). *Les Structures Elémentaires de la Parenté*, París-La Haya: Mouton. [Versión castellana: *Estructuras elementales del parentesco.* Barcelona, Planeta-Agostini, 1ª ed., 1985.]
Lévi-Strauss, C. (1967-71). *Mythologiques*, I-IV, París:Plon
Lévi-Strauss, C. (1973). *Anthropologie Structurale Deux*, París: Plon.
Lewis, C.I. y Langford, C.H. (1932). *Symbolic Logic*, Nueva York: Dover.
Lewis, D. (1968). "Counterpart theory and quantified modal logic", *Journal of Philosophy,* 65, 113-126.
Lewis, D. (1973a). "Causation", *Journal of Philosophy*, 70, 556-567.
Lewis, D. (1973b). *Counterfactuals*, Oxford: Blackwell.
Lindbeck, A. (1976). "Stabilization policy in open economies with endogenous politicians", *American Economic Review: Papers and Proceedings,* 66, 1-19.
Lipset, S.M. y Bendix, R. (1959). *Social Mobility in Industrial Society*, Berkeley y Los Angeles: University of California Press.
Lipset, S.M. y Gordon, J. (1953), "Mobility and trade union membership", en S.M.Lipset y R.Bendix (comps.), *Class, Status and Power*, Glencoe, Ill: Free Press.
Loschky, D.J. (1973). "Studies of the Navigation Acts: New economic non-history", *Economic History Review*, 26, 689- 691.
Luce, R.D. y Raiffa, H. (1957). *Games and Decisions*, Nueva York: Wiley.
Luxemburgo, R. (1899). "Sozialreform order Revolution?" en *Gesammelte Werke,* vol.1/1, Berlin:Dietz, 1970. [Versión castellana: *Reforma o Revolución.* Barcelona, Grijalbo, 1ª ed., 1974.]
Luxemburgo, R. (1906). "Massenstreik, Partei und Gewerkschaften", en *Gesammelte Werke,* vol.2, Berlin:Dietz, 1972. [Versión castellana: *Huelga de masas, partidos y sindicatos.* Barcelona, Grijalbo, 1ª ed., 1975.]
Lyons, D. (1965). *Forms and Limits of Utilitarianism,* Oxford: Oxford University Press.
Maarek, E. (1975). *Introduction au Capital de Karl Marx*, París: Calmann-Lévy.
McClelland, P.D. (1975). Causal Explanation and Model- Building in *History, Economics and the New Economic History*, Ithaca, NY: Cornell University Press.
McFarland, D. (1970). "Intergenerational social mobility as a Markov process", *American Sociological Review,* 35, 463- 476.
McFarland, D. (1976). "Putting goal-oriented actors into Harrison White's vacancy chains" trabajo presentado ante el Annual Meeting of the American Sociological Association, Nueva York.
McGuire, W.J. (1966)."The status of cognitive consistency theories", en S.Feldman (comp.), *Cognitive Consistency*, Nueva York: Academic Press.
MacIntyre, A. (1962). "A mistake about causality in the social sciences", en P.Laslett y W.C.Runciman (comps.), *Politics, Philosophy and Society*, 2a. serie, Oxford: Blackwell.
Mackie, J.L. (1965). "Causes and conditions", *American Philosophical Quarterly*, 2, 245-264.
Mackie, J.L. (1967). "Fallacies", *The Encyclopedia of Philosophy*, Nueva York: Macmillan.
Mandeville, B. (1729). *The Fable of the Bees*, 6a. edición, Londres.
Manning, P. (1974). "Analysing the costs and benefits of colonialism", *African Economic History Review*, 1, 15-22.
Marx, K. (1845-46). "Die Deutsche Ideologie, en" *Marx-Engels Werke*, vol.3, Berlín: Dietz. [Versión castellana: *La ideología alemana,* Barcelona, Grijalbo, 1970.]

Marx, K. (1847). "Die moralisierende Kritik und die Kritisierende Moral", en *Marx-Engels Werke*. vol.4, Berlin:Dietz.
Marx, K. (1850). *The Class Struggles in France*, citado de la edición Nueva York: International Publishers, 1935. [Versión castellana: *Lucha de clases; El dieciocho brumario*. Madrid, Espasa-Calpe, 2ª ed., 1992.]
Marx, K. (1857). "The secret diplomatic history of the 18th- century", citado de la reedición de R.Payne (comp.), *The Unknown Karl Marx*, Nueva York: New York University Press.
Marx, K. (1857-58). "Grundrisse der Kritik der politischen Ökonomie," citado de la traducción de Pelican, Londres 1973. [Versión castellana: *Elementos fundamentales para la crítica de la economía política*. Madrid, Siglo XXI, 2ª ed., 1976].
Marx, K. (1862-63). "Theorien Über den Mehrwert", en *Marx-Engels Werke*, vol.26, Berlín: Dietz.
Marx, K. (1867). *Capital I*, Nueva York: International Publishers.
Marx, K. (1871). *The Civil War in France*, citado de la edición de Moscú: Foreign Languages Publishing House, 1952. [Versión castellana: *La guerra civil en Francia*. Madrid, Aguilera, 4ª ed., 1976.]
Marx, K. (1879-80). "Randglossen zu Adolph Wagners 'Lehrbuch der politischen Ökonomie'", en *Marx-Engels Werke*, vol. 19, Berlín: Dietz.
Marx, K. (1894). *Capital III*, Nueva York: International Publishers.
Marx, K. (1933). "Resultate des unmittelbaren Produktionsprozesses", *Arkhiv Marksa i Engelsa*, II (VII).
Mates, B. (1968). "Leibniz on possible worlds", en B.van Rootselaar y J.F.Staal (comp.), *Logic, Methodology and Philosophy of Science III*, Amsterdam: North-Holland.
Meisner, M. (1967). *Li Ta-chao and the Origins of Chinese Communism*, Cambridge, Mass: Harvard University Press.
Merleau-Ponty, M. (1945). *Phénomenologie de la Perception*, París: Gallimard. [Versión castellana: *Fenomenología de la percepción*. Barcelona, Edicions 62, 2ª ed., 1980.]
Merleau-Ponty, M.(1963). *La Structure du Comportement*, París: Presses Universitaires de France.
Merton, R. (1936). "The unanticipated consequences of purposive social action", *American Sociological Review, 1*, 894-904.
Merton, R. (1957). *Social Theory and Social Structure*, Glencoe, Ill.:Free Press.
Merton R. (1973). *The Sociology of Science*, Chicago: Chicago University Press. [Versión castellana: *Sociología de la ciencia*. Madrid, Alianza Editorial, 2ª ed., 1985.]
Milgram, S. (1974). *Obedience to Authority*, Londres: Tavistock. [Versión castellana: *Obediencia a la autoridad*. Bilbao, Editorial Española Desclee de Brouwer, 3ª ed., 1984.]
Mill, J.S. (1967). *A System of Logic*, Londres: Longman.
Mondadori, F. (1973). "Reference, essentialism and modality in Leibniz's metaphysics", *Studia Leibnitiana*, 5, 74-101.
Moore, G.E. (1922). "External and internal relations", en G.E.Moore, *Philosophical Studies*, Londres: Rutledge.
Morishima, M. (1973). *Marx's Economics*, Oxford: Oxford University Press. [Versión castellana: *La teoría económica de Marx*. Madrid, Tecnos, 1ª ed., 1976.]
Mortimore, G.W. (comp.) (1971). *Weakness of Will*, Londres: Macmillan.
Mukherjee, R. (1954). "A study of mobility between three generations", en D.Glass (comp.), *Social Mobility in Britain*, Londres: Routledge.
Nagel, E. (1963). "Assumptions in economic theory", *American Economic Review, 53*, 211-236.
Nathan, A. (1976). "Policy oscillations in the People's Republic of China: A critique", *China Quarterly*, 68, 720-733.
Needham, R. (1962). *Structure and Sentiment*, Chicago: Chicago University Press.
Neher, P. (1971). "Peasants, procreation and pensions", *American Economic Review*, 61, 380-389.
Nordhaus, W. (1975). "The political business cycle", *Review of Economic Studies*, 42, 169-190.
North, D. (1971). "Institutional change and economic growth", *Journal of Economic History*.

North, D. y Davis, L. (1971). *Institutional Change and American Economic Growth,* Cambridge: Cambridge University Press.
North, D. y Thomas, R.P. (1973). *The Rise of the Western World,* Cambridge: Cambridge University Press. [Versión castellana: *El nacimiento del mundo occidental.* Madrid, Siglo XXI, 5ª ed., 1991.]
Nozick, R. (1969). "Newcomb's problem and two principles of choice", en N.Rescher (comp.), *Essays in Honor of Carl G. Hempel,* Dordrecht: Reidel.
Nozick, R. (1974). *Anarchy, State and Utopia,* Oxford: Blackwell.
Nuti, D.M. (1970). "Capitalism, socialism and steady growth", *Economic Journal,* 80, 32-57.
Nutini, H.G. (1970). "Some considerations on the nature of social structure and model building", en E.N.Hayes y T. Hayes (comps.), *Claude Lévi-Strauss: The Anthropologist as Hero,* Cambridge, Mass: MIT Press.
Okun, A. (1975). *Equality and Efficiency: The Big Trade-Off,* Washington: The Brookings Institution.
Olson, M. (1963). *The Logic of Collective Action,* Cambridge, Mass: Harvard University Press.
Ollman, B. (1976). *Alienation: Marx's Concept of Man in Capitalist Society,* 2a. edición, Cambridge: Cambridge University Press. [Versión castellana: *Alienación.* Buenos Aires, Amorrortu Editores, 1ª ed., 1975.]
Owen, G. (1968). *Game Theory,* Filadelfia: Saunders.
Parsons, T. (1949). *The Structure of Social Action,* Glencoe, Ill.:Free Press. [Versión castellana: *La estructura de la acción social.* Barcelona, Guadarrama.]
Pen, A. (1959). *The Wage Rate under Collective Bargaining,* Cambridge, Mass.: Harvard University Press.
Perkins, H. (1969). *The Origins of Modern English Society: 1780-1880,* Londres: Routledge.
Peyrefitte, A. (1976). *Le Mal Français,* París: Plon. [Versión castellna: *El mal latino.* Barcelona, Plaza y Janés Editores, 1ª ed., 1980.]
Phelps, E.S.y Pollak, R.A. (1968). "On second-best national saving and game-equilibrium growth", *Review of Economic Studies,* 35, 185-199.
Pollak, R.A. (1968). "Consistent planning", *Review of Economic Studies,* 35, 201-208.
Pollak,R.A. (1976a). "Habit-formation and long-run utility functions", *Journal of Economic Theory,* 13, 272-297.
Pollak, R.A. (1976b). "Interdependent preferences", *American Economic Review,* 66, 309-320.
Popper, K. (1940). "What is dialectic?", *Mind,* 49, 403-426.
Popper, K. (1963). *Conjectures and Refutations,* Londres: Routledge. [Versión castellana: *Conjeturas y refutaciones.* Barcelona, Paidós, 3ª ed., 1991.]
Pörn, I. (1970). *The Logic of Power,* Oxford: Blackwell.
Portal, R. (1965). "The industrialization of Russia", en H.J. Habakkuk y M.Postan (comps.), *Cambridge Economic History of Europe,* Cambridge: Cambridge University Press.
Postal, P.M. (1964). "Limitations of phrase structura grammars", en J.A.Fodor y J.J.Katz (comps.), *The Structure of Language,* Englewood Cliffs, N.J.:Prentice-Hall.
Prior, A.N. (1967). *Past, Present and Future,* Oxford: Oxford University Press.
Quine, W.V.O. (1953). *From a Logical Point of View,* Cambridge, Mass: Harvard University Press. [Versión castellana: *Desde un punto de vista lógico.* Barcelona, Ariel, 1962.]
Quine, W.V.O. (1959). *Methods of Logic,* Nueva York: Holt- Dryden. [Versión castellana: *Los métodos de la lógica.* Barcelona, Ariel, 2ª ed., 1967.]
Quine, W.V.O. (1960). *Word and Object.* Nueva York: Wiley. [Versión castellana: *Palabra y objeto.* Barcelona, Labor, 1968.]
Quirk, J. y Saposnik, R. (1968). *Introduction to General Equilibrium Theory and Welfare Economics,* Nueva York: McGraw-Hill. [Versión castellana: *La teoría del equilibrio general y la economía del bienestar.* Barcelona, Bosch Casa Editorial.]
Rabin, M. y Scott,D. (1959). "Finite automata and their decision problems", *IBM Journal of Research and Development, 3,* 114-125.
Rader, T. (1971). *The Economics of Feudalism,* Nueva York: Gordon and Breach.
Raiffa, H. (1968). *Decision Analysis,* Conferencia, Mass.: Addison-Wesley. [Versión castellana: *Análisis de la decisión empresarial.* Bilbao, Ediciones Deusto, 1ª ed., 1977.]

Rapoport, A. (1966). *Two-Person Game Theory*, Ann Arbor: University of Michigan Press.
Rapoport, A. (1975). "Comment on Bram's discussion of Newcomb's problem", *Journal of Conflict Resolution*, 19, 613-619.
Rapoport, A. y Chammah, A. (1965), *Prisoner's Dilemma*, Ann Arbor: University of Michigan Press.
Rapoport, A., Guyer, M. y Gordon, D. (1976). *The 2 x 2 Game*, Ann Arbor: University of Michigan Press.
Rawls, J. (1971). *A Theory of Justice*, Cambridge, Mass: Harvard University Press. [Versión castellana: *Teoría de la justicia*. Madrid, Fondo de Cultura Económica, 1ª ed., 1979.]
Rescher, N. (1967). *The Philosophy of Leibniz*, Englewood Cliffs, NJ: Prentice-Hall.
Rescher, N y Urquhar, A.(1971). *Temporal Logic*, Viena-Nueva York: Springer.
Riker, W. y Ordeshook, P. C. (1973) *An Introduction to Positive Political Theory*, Englewood Cliffs, NJ: Prentice-Hall.
Robinson, J. (1956). *The Accumulation of Capital*, London: Macmillan.
Rorty, R. (1967). "External and internal relations", *The Encyclopedia of Philosophy*, Nueva York: Macmillan.
Rosenberg, N. (1974). "Science, invention and economic growth", *Economic Journal*, 84, 90-108.
Russel, R. (1941). "The effects of slavery upon nonslaveholders in the ante-bellum South", en H.D.Woodman (comp.), *Slavery and the Southern Economy*, Nueva York: Pantheon.
Russell, B. (1905). "On denoting", *Mind*, 14, 479-493.
Russell, B. (1910). "Some explanation in reply to Mr. Bradley". *Mind*, 19, 373-378.
Russell, B. (1938). *Power*, Londres y Nueva York: Norton.
Salter, W. (1960). *Productivity and Technical Change*, Cambridge: Cambridge University Press. [Versión castellana: *Productividad y cambio técnico*. Madrid, Ministerio de Trabajo y Seguridad Social. Centro de Publicaciones, 1ª ed., 1986.]
Salthe, S.N. (1972). *Evolutionary Biology*, Nueva York: Holt, Rinehart and Winston.
Samuelson, P. (1950). "The evaluation of real national income", *Oxford Economic Papers*, 2, 1-29.
Samuelson, P. (1957). "Wages and interest", *American Economic Review*, 48, 884-912.
Sartre, J.-P. (1943). *L'Etre et le Néant*, París: Gallimard. [Versión castellana: *El ser y la nada*. Madrid, Alianza Editorial, 2ª ed., 1989.]
Sartre, J.-P. (1960). *Critique de la Raison Dialectique*, París: Gallimard.
Schelling, T.S. (1971). "Dynamic models of segregation", *Journal of Mathematical Sociology*, 1, 143-186.
Schmookler, J. (1966). *Inventions and Economic Growth*, Cambridge, Mass: Harvard University Press.
Schumpeter, J.A. (1951). *Imperialism and Social Classes*, Nueva York: Kelley. [Versión castellana: *Imperialismo y clases sociales*. Madrid, Tecnos, 1965.]
Schumpeter, J.A. (1953). *Capitalism, Socialism and Democracy*, Londres: Allen and Unwin. [Versión castellana: *Capitalismo, socialismo y democracia*. Madrid, Aguilar, 1971.]
Schumpeter, J.A. (1954). *A History of Economic Analysis*, Londres: Allen and Unwin. [Versión castellana: *Historia del análisis económico*. Barcelona, Ariel, 2ª ed., 1982.]
Schurmann, F. (1967). *Ideology and Organization in communist China*, Berkeley y Los Angeles: University of California Press.
Scitovsky, T. (1971). *Welfare and Competition*, Londres: Allen and Unwin. [Versión castellana: *Bienestar y competencia*. Buenos Aires, Amorrortu Editores, 1ª ed., 1967.]
Seip, J.A. (1958). *Teorien om det Opinionsstyrte Enevelde*, Oslo: Universitetsforlaget.
Sejersted, F. (1975), "En Teori om Embedsmannsstaten", Oslo, Historisk Institutt (mimeo).
Selden, M. (1971). *The Yenan Way in Revolutionary China*, Cambridge, Mass: Harvard University Press.
Sen, A.K. (1967). "Isolation, assurance and the social rate of discount", *Quarterly Journal of Economics*, 80, 112- 124.
Sen, A.K. (1970). *Collective Choice and Social Welfare*, San Francisco: Holden-Day. [Versión

castellana: *Elección colectiva y bienestar social.* Madrid, Alianza Editorial, 1976.]
Sen, A.K. (1973). *On Economic Inequality,* Oxford: Oxford University Press. [Versión castellana: *Sobre la desigualdad económica.* Barcelona, Crítica, 1979.]
Sen, A.K. (1976). "Real national income", *Review of Economic Studies,* 42, 19-39.
Shubik, M. (1970). "Game theory, behavior and the paradox of the prisoner's dilemma", *Journal of Conflict Resolution,* 14, 181-202.
Simensen, J. (1977). "Counterfactual arguments in historical analysis". Del debate sobre la partición de Africa y el efecto del dominio colonial (mimeo).
Simmel, G. (1908). *Soziologie,* Berlín: Duncker und Humboldt. [Versión castellana: *Sociología.* Madrid, Alianza Editorial, 1986.]
Simon, H. (1953). "Notes on the observation and measurement of political power", *Journal of Politics,* 15, 500-516.
Simon, H. (1954a). "Bandwagon and underdog effects in election predictions", *Public Opinion Quarterly,* 18, 245-253.
Simon, H. (1954b). "A behavioral theory of rational choice", *Quarterly Journal of Economics,* 69, 99-118.
Singer, J.E. (1966). "Motivation for consistency", en S.Feldman (comp.), *Cognitive Consistency,* Nueva York: Academic Press.
Singer, M. (1963). *Generalization in Ethics,* Londres: Eyre and Spottiswoode.
Skinner, G.W. y Winckler, E.A. (1969). "Compliance succession in rural Communist China: A cyclical theory", en A.Etzioni (comp.), *A Sociological Reader on Complex Organizations,* Nueva York:Holt, Rinehart and Winston.
Smelser, N. (1959). *Social Change in the Industrial Revolution,* Londres: Routledge.
Smith, J. (1968). "Transportation", en *The International Encyclopedia of the Social Sciences,* Nueva York: Macmillan.
Snyder, D.P. (1971). *Modal Logic and its Aplications,* Nueva York: van Nostrand, Reinhold.
Sobel, J.H. (1967). "'Everyone', consequences and generalization arguments", *Inquiry,* 10, 373-404.
Sosa, E. (1975). Introducción a E.Sosa (comp.), *Causation and Conditionals,* Oxford: Oxford University Press.
Staaf, R. y Tannian, F. (comps.) (1972). *Externalities,* Nueva York: Dunnellen.
Stalnaker, R. (1968). "A theory of conditionals", en N.Rescher (comp.), *Studies in Logical Theory,* Oxford: Blackwell.
Stampp, K. (1976). "*Introducción*" a P.David y cols., *Reckoning with Slavery,* Nueva York: Oxford University Press. [Del mismo autor: *La esclavitud en los E.E.U.U.* Vilassar de Mar, Oikos-Tau, 1966.]
Stinchcombe, A. (1968). *Constructing Social Theories,* Nueva York: Harcourt, Brace and World.
Stouffer, S.A. y cols. (1949). *The American Soldier,* Princeton: Princeton University Press.
Strotz, R.H. (1955-56). "Myopia and inconsistency in dynamic utility maximization", *Review of Economic Studies,* 23, 165-180.
Suttmeier, R. (1974). *Research and Revolution,* Lexington, Mass.: Lexington Books.
Sweezy, P. (1962). *The Theory of Capitalist Development,* Londres: Dennis Dobson. [Versión castellana: *Teoría del desarrollo capitalista.* Madrid, Fondo de Cultura Económica, 10ª ed., 1977.].
Takayama, A. (1974). *Mathematical Economics,* Hinsdale, Ill.: The Dryden Press.
Taylor, A.J. (1960). "Progress and poverty in Britain 1750- 1850", *History,* 45, 16-31.
Taylor, C. (1971). "Interpretation and the sciences of man", *Review of Metaphysics,* 25, 3-51.
Taylor, M. (1976). *Anarchy and Cooperation,* London: Wiley.
Taylor, M. (1977). "Altruism and cooperation: Their evolution by natural selection and their role in primitive societies" (mimeo).
Thalheimer, A. (1930). "Über den Faschismus", en W.Abendroth (comp.), *Faschismus und Kapitalismus,* Francfort-Viena: Europäische Verlagsanstalt 1967.
Thernstrom, S. (1970). "Working class social mobility in industrial America", en M.H.Richter (comp.), *Essays in Theory and History,* Cambridge, Mass.: Harvard University Press.

Thernstrom, S. (1973). *The Other Bostonians*, Cambridge, Mass.: Harvard University Press.
Thomas, R.P. (1965). "A quantitative approach to the study of the effects of British imperial policy upon colonial welfare", *Journal of Economic History*, 25, 615-638.
Thompson, E.P. (1968). *The Making of the English Working Class*, Harmondsworth: Pelican Books.
Thompson, E.P. (1971). "The moral economy of the English crowd in the 18th century", *Past and Present*, 50, 76- 136.
Trotsky, L. (1972). *1905*, Londres: Allen Lane.
Tullock, G. (1974). "Does punishment deter crime?", *The Public Interest*, 36, 103-111.
Ullman-Margalit, E. (1976). "The generalization argument", *Journal of Philosophy*, 73, 511-522.
Vendler, Z. (1967). "Any and all", *The Encyclopedia of Philosophy*, Nueva York: Macmillan.
Vernon, R. (1971). *Sovereignty at Bay*, Londres: Longman.
Vlastos, G. (1967). "Zeno of Elea", *The Encyclopedia of Philosophy*, Nueva York: Macmillan.
Walton, G. (1971). "The economic history and the burdens of the Navigation Acts", *Economic History Review*, 24, 533- 542.
Watzlawick, P.H., Beavin, J. y Jackson, D. (1967). *Pragmatics of Human Communication*, Nueva York: Norton. [Versión castellana: *Teoría de la comunicación humana*. Barcelona, Herder, 8ª ed., 1991.]
Weber, M. (1968). "Objektive Möglichkeit and adäquate Verursachung in der historischen Kausalbetrachtung" en M.Weber, *Gesammelte Aufsätze zur Wissenschaftlehre*, Tubinga: Mohr.
Weil, E. (1966). *Hegel et l'Etat*, París: Vrin.
Weiszacker, C. C. von (1971a), Notes on endogenous change of tastes, *Journal of Economic Theory*, 3, 345-372.
Weiszäcker, C.C. von (1971b). *Steady-State Capital Theory*, Nueva York: Springer.
Whitrow, G. (1961). *The Natural Philosophy of Time*, Edimburgo: Nelson.
Wildavsky, A. (1964). *The Politics of the Budgeting Process*, Boston: Little, Brown.
Wilensky, E. y Edwards, H. (1959). "The Skidder", *American Sociological Review*, 24, 215-231.
Williams, G.C. (1966). *Adaptation and Natural Selection*, Princeton: Princeton University Press.
Winch, P. (1958). *The Idea of a Social Science*, Londres: Routledge.
Winckler, E. (1976). "Policy oscillations in the People's Republic of China: A reply", *China Quarterly*, 68, 734- 750.
Winter, S.G. (1964). "Economic 'Natural selection' and the theory of the firm", *Yale Economic Essays*, 4, 225-272.
Winter, S.G. (1975). "Optimization and evolution", en R.H.Day y T. Groves (comps.), *Adaptative Economic Models*, Nueva York: Academic Press.
Wittgenstein, L. (1953). *Philosophische Untersuchungen*, Oxford: Blackwell. [Versión castellana: *Investigaciones filosóficas*. Barcelona, Crítica, 1988.]
Yasuba, Y. (1961). "The profitability and viability of plantation slavery in the United States", *The Economic Quarterly Studies Review*, 12, 60-67.
Young, G. (1976). "The fundamental contradiction of capitalist production", *Philosophy and Public Affairs*, 5, 196-234.

Indice temático

Alienación, 143
Análisis marginal, 142-143, 250
Aprendizaje,83, 107, 176
Aristóteles, 72, 79
Arrow. K., 171, 179, 185, 193

Barcan, fórmula de, 48
Bateson, G., 112, 176
Becker, G., 204, 205
Bernstein, E., 186
Bienestar,
 promedio contra agregado, 245, 246
 ingreso nacional real, 67-71, 242, 262
Bossuet, 144, 145
Boudon, R., 164, 165, 166, 167, 189, 205

Caída de la tasa de ganancia, 152, 158
Capitalismo, 134, 143, 172-176, 186, 190, 191, 247
Cártel, 165, 170, 171
Cambio,
 en las preferencias, 71-75, 114, 207
 social, 175, 194
 técnico, 251-254
Causalidad, 72, 116-118, 201-208, 219-220, 233-236, 242
Climo, T.A., 234-235
Composibilidad, 221
Contradicciones, 100-193
 y cambio, 31-32, 103-107, 127, 175-178, 193-194
 y dialéctica, 31-32, 101-107, 146
 en Marx, 134, 141, 152-158, 171-173
Contrafácticos, 219-268
 y condicionales, 42
 condiciones de afirmabilidad, 225, 226, 237
 condiciones de legitimidad, 229, 230, 236, 237, 247-249
 teoría metalingüística de, 226-227
 teoría ontológica de, 222, 226, 232-234
 condiciones de verdad, 226, 227
Contrafinalidad, 90, 144,-158
Contramovilidad, 189
Contrapartes, 222, 238
Creencia,
 contradictoria, 106, 107, 118, 128
 racional, 121, 122
 revelada, 121, 124
Crisis, 134, 157, 192
Cuantificador, 42

David, P., 153, 155
Davidson, D., 44, 117, 118
Davis, L., 179, 193
Debilidad de la voluntad, 117
Delito, 205, 206
Determinismo, 44, 81, 222-225, 228
Dialéctica, 31, 94-96, 100-103
Dickinson, E., 29, 95, 112, 130
Disonancia cognitiva, 123-127
Doble vínculo 85, 112
Donne, J. , 52, 53, 95, 108, 109, 110, 129
Dummett, M., 53, 88, 222, 238
Duopolio, 152

Educación, 165-167
Engerman, S., 242, 257-258, 263-266
Esclavitud, 106-111, 256-267
Estado, el, 161, 162, 190, 191
Estrategia, 194, 195
 dominante, 163, 166, 197
 maximin, 167, 197
 mixta, 164, 167, 168, 196
Estructuralismo, 64-67, 204
Explotación, 31, 52, 142, 143, 174, 214, 248

Falacias,
 ecológica, 137
 de composición, 49, 135-144, 250
 de división, 135-137
 patética, 104, 157
 estructuralista, 158-161
Fellner, W., 153
Festinger, L., 123, 124
Fetichismo, 56
Finley, M.I., 149, 213
Fogel, R. 34, 251-258, 261-266
Forrester, J., 148
Freud, S., 112
Frustraciones relativas, 167, 189
Funcionalismo, 161, 162, 186, 190

Genovese, E., 108, 110, 111, 149, 210, 261, 265

Gerschenkron, A., 149, 173
Goldman, A., 82, 230
Gramáticas
independientes del contexto, 75-77
de estados finitos, 62-64
Granovetter, M., 184-193

Hartwell, R.M., 242, 245, 246
Hayek, F., 247
Hegel, G.W.F., 31, 96, 100-113, 132, 149, 209, 265, 266
Hintikka, J., 53, 57, 110, 117-123, 209
Hirschman, A.O., 87, 179, 213
Histéresis, 81, 97
Hobsbawm, E., 159, 244-246
Howells, P.G.A., 234-235
Hume, D., 139-140
Hylland, A., 80, 98

Identidad, 222, 223
Imperialismo, 170, 213, 239
Inercia acumulativa, 61, 87
Información, 125-127, 164, 197
Intencionalidad, 72, 82-85, 158, 202-203
Interpretación, 41, 45, 48
Irracionalidad, 113-117, 126, 264-265
Irreversibilidad, 94-96, 192, 255

Juegos,
de certeza, 163, 164, 197, 199, 203, 264
de la gallina, 153-197-198
sin estrategias dominantes, 164, 167, 170, 200
sin puntos de equilibrio, 177, 195
sin solución, 153, 177, 178, 195, 197, 265

Kant, I., 140, 208
Kenny, A., 79, 125, 142
Keynes, J.M., 134, 145
Kripke, S., 53, 88, 222, 238

Lancaster, K., 174, 175
Leibniz, G.W., 46, 47, 53, 54, 90, 98, 109 213, 221
Lévi-Strauss, C., 64-67
Lenin, N., 93, 187
Lewis, D., 222, 225-240, 251, 255, 259, 260, 268
Libertad, 202, 203, 206
Lingüística, 63-66, 77, 83

Lógica,
de la cuantificación, 42-45
deóntica, 47, 138
epistémica, 47, 121
modal, 44-51
proposicional, 39-43
temporal, 31, 47
Luxemburgo, R., 93, 186

MacIntyre, A., 118
Malebranche, N., 145
Mandeville, B., 145
Mano invisible, 144-146, 161
Marx, K., 31, 32, 50, 51, 92-95, 134, 146, 152-158, 171-174, 177, 181-182, 186, 192, 207-215
Merton, R,, 161
Método comparativo, 220, 267-269
Mill, J.S., 223,
Modalidades iteradas, 45-46, 86
Modelo de la telaraña, 148, 151, 163-165, 176
Movilidad, 59-62, 184-192
Mondadori, F., 233
Moore, G.E., 53
Mukherjee, R., 190
Mundos posibles 35, 46, 47, 81, 86, 87, 96, 97, 141, 221, 226, 231, 232, 236

Nash, J., 208
Necesidad, 45-46, 60-62, 85-89
Negación de la negación, 94-96
Newcomb, problema de, 122-123
No decisiones, 191
Normas, 204-205
North, D., 179-180, 193, 218
Nozick, R., 122, 144, 146, 160, 208

Olson, 164-165
Organización, 164-166, 179-194

Poder, 51, 53-55, 82-84, 108, 175, 230-232
Popper, K., 100, 103, 105, 162
Posibilidad, 36, 37, 45, 60-70, 86, 223
Preferencias,
cambiantes, 71-75, 114, 206
consistentes, 113, 114, 131
temporales, 114-116, 168-169
Prisionero, Dilema del, 162-178, 196-200
Propiedades esenciales, 52, 268

Quine, W.V.O., 41, 107

Radicalismo, 187, 188
Rapoport, A., 123, 199
Racionalidad, 119, 122, 123, 147, 150, 153, 158, 177, 196, 204-206, 218, 264
Rawls, J., 243
Reconocimiento, 108, 109, 126
Reflexividad, 46, 91
Relaciones, 42, 43, 50-56
Rousseau, J.-J., 109
Russel, R., 260, 261
Russell, B., 49, 82, 125

Samuelson, P., 67-70
Sartre, J.-P., 111, 112, 133, 148
Schelling, T.S., 148
Schmookler, J., 252, 253
Schumpeter, J., 171
Sejersted, F., 169
Selden, M., 182, 183
Sen, A.K., 70, 71
Simensen, J., 238-240
Simmel, G., 182, 183
Simetría, 46, 94-96
Simon, H., 149, 206
Sindicatos, 165, 182, 184, 187
Sistemas de parentesco, 66, 67
Smelser, N., 153
Smith, A., 144, 145
Solidaridad, 143, 180, 181, 187
Solución, 195
Stalnaker, R., 225, 226, 227, 231
Stinchcombe, A., 161
Strotz, R.H., 115
Suboptimalidad, 162-175
Sweezy, P., 148, 211

Taylor, M., 200
Thernstrom, S., 188
Thompson, E.P., 159, 186
Tiempo,
  en la política, 86, 87
  y explotación, 143, 174
  y organización, 188, 189, 193
  y preferencias, 71-75, 114-116, 167, 169
  y suboptimalidad, 168-175
Transitividad, 46, 91-94
Trotsky, 93

Ullman-Margalit, E., 137, 140
Universalizabilidad, 137, 138, 139
Utopía, 91, 92, 173

Weber, M., 224, 225
Weiszäcker, C.C. von, 72, 73, 74, 114
Williams, G.C., 160
Winch, P., 139, 140
Winter, S., 206
Wittgenstein, L., 139, 140

FILOSOFIA

(viene de la pág. 4)

PIER ALDO ROVATTI	*Como la luz tenue*
GEORGES BALANDIER	*El desorden*
HANNAH ARENDT	*Hombres en tiempos de oscuridad*
PAUL RICOEUR	*Ideología y utopía*
ERNEST GELLNER	*Cultura, identidad y política*
JEAN FRANÇOIS LYOTARD	*La diferencia*
RONALD DWORKIN	*El imperio de la justicia*
CORNELIUS CASTORIADIS	*Los dominios del hombre*
KARL JASPERS	*La práctica médica en la era tecnológica*
LUDWIG WITTGENSTEIN	*Sobre la certeza*
JEAN FRANÇOIS LYOTARD	*El entusiasmo*
JEAN FRANÇOIS LYOTARD	*La posmodernidad (explicada a los niños)*
TZVETAN TODOROV	*Frágil felicidad*
GIANNI VATTIMO	*El fin de la modernidad*
GIANNI VATTIMO	*Introducción a Heidegger*
JACOB BRONOWSKI	*Los orígenes del conocimiento y la imaginación*
NORBERTO BOBBIO	*El problema de la guerra y las vías de la paz*
GILLES DELUZE	*Empirismo y subjetividad*